

Oy	666	AGAGCGACGGAGTCCGCGCGAAGCGGCTTCGCGTGGCCGCGACGCGCGAAGCGAAGACA	725
Db	855	AGGACGGGGAGAGAGACGAGGACCGGGGAGGACGGGAGAGAGACGAGGACGGGAGGACG	796
Oy	726	ACGGCCATGGCGTGGCTGCCCGCTGAAACGGCAAGAGCAAGCGCCATGGCGTGGTGC	785
Db	795	GGGAGGACGAGGACGGGGAGGACGGGGAGGACGAGACGGGGAGGACGGGAGGAGCGAGG	736
Oy	786	ACCGGAACCGCGAAGACGACCGCCATGGCGCTGGCTGCCGAGCGGACCGCGAAGGACACG	845
Db	735	ACGGGGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGGGG	676
Oy	846	GCCATGCGGAGGCGCATCTGCGAACGGCGACCGCGAGGCCACTGCGAACCGCGAAGACCA---	902
Db	675	AGGACGGGGAGGAGGACGAGGACGGGGAGGACGAGGACGGGAGGACGGGGAGGACGGGG	616
Oy	903	ACGGCCACCGCGAGAGCACAACGGCCATGCTGAGGCCCGCGACGCGAACCGCGAGGACCAAC	962
Db	615	AGGACGAGGACCGGGAGGAGGACGAGGACGGGGAGGACCGGGAGGACCGGGAGGAGGACG	556
Oy	963	AGCATGCGGAGGACTTCGCGCGCGAAGCGCGAGCAACCGGCGCTATGGCGCGCGCGCGACG	1022
Db	555	AGGACGGGGAGGACCGGGAGGAGGACGAGGACGGGGAGGAGGACGAGGACGGGGAGGACG	496

Qy	1023	AGGAGAGAGAGGCGGTGGATGGAAATTTCCGGGGTGGCCAAAGACCGGGCTGTCGGGGA	1082
Db	495	GGGAGGACCGGGGAGGAGGACGAGACCGGGGAGACCGGGAGAGGACGAGACCGGGAGG	436
Qy	1083	CGGGGCGGAACTATGACATCCGGGCGATACGGTACAAAGTACGGCGGAGCTGCAGGAGA	1142
Db	435	AGGACGAGAGACCGGGAGGACCGGGGAGGACCGGGGAGAGAGACGAGGACCGGGAGGAGCG	376
Qy	1143	AGGGGCGCGGCGCCCTGTCTGCGCGCTGGCCACCGGGAGACCGTCCGTGTTCCCGGCTTTCC	1202
Db	375	AGGACCGGGGAGGACGGGGAGGACCGGGGAGGACCGGGGAGGAGGACGAGGACGGGAGGAGG	316
Qy	1203	GCAAGCGCGTGCAGAGCGGAGGACCGCGGCGGCG	1235
Db	315	ACGAGGACCGGGAGGACGAGGACCGGGAGGGCG	283

```

RESULT 2
US-09-130-114-2/c
Sequence 2, Application US/09130114
Patent No. 5976807
GENERAL INFORMATION:
APPLICANT: HORLICK, ROBERT A.
APPLICANT: DAMAJ, BESSAM B.
APPLICANT: ROBBINS, ALAN K.
TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
FILE REFERENCE: 0867/1D9030U1
CURRENT APPLICATION NUMBER: US/09/130,114
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1931
TYPE: DNA
ORGANISM: EBNA
US-09-130-114-2

```

Query Match	0.9%	Score 96.2;	DB 2;	Length 1931;
Best Local Similarity	49.2%;	Pred. No. 1.1e-11;		
Matches 282; Conservative	0;	Mismatches 288;	Indels 3;	Gaps 1.

QY	666	AGAGCGACGGAAGTCCGCCCGCAACGGCCCTTGCCGTGGCCGACGCCCGCAACGGCAAGAGCA	725
Db	855	AGGACGGGGAGAGAGACGAGGACGGGGAGGACGGGGAGAGAGAGAGAGAGCAACGGGGAGGACG	796
QY	726	ACGGCCATGCGCTGTGCTCGCGCCCTGCAACGGCAAGACCAACGGCCATGTGCGCTGATCCG	785
Db	795	GGGAGGACGAGGACCGGGGAGGACGGGGAGGACGAGGACGGGAGAGACGGGGAGGACGAGG	736
QY	786	ACCGGAAACGGCMAAGACAAACGGCCATGGCTGTGCTGCCGACCGCAACGGCMAAGACAAACG	845
Db	735	ACGGGAGGAGCAGAGGACGGGGAGGACGAGGACGGGGAGGACGAGAGGACGGGGAGGACCGGG	676
QY	846	GCCATGCGGAGGCGCACTGTGCGAACGGGCCGAGGCGCACTTGCAACGGCMAAGCA---	902
Db	675	AGGACGGGGAGAGAGAGACGAGGACGGGGAGGACGAGGACGGGGAGGACGGGGAGGACGGGG	616
QY	903	ACGGCCACCGCGAGAGCAACGGCCATGAGCGCGCCGACCGCAACGGCGAGAGCAACG	962
Db	615	AGGACGAGGACCGGGAGGAGGACGAGGACGGGGAGGACGGGGAGGACGGGGAGGAGGACG	556
QY	963	AGCATGCGGAGGACTCCGGCGCGAAACGGCGAGACCAACCGGCATGCGCGCGCGCGCGCAG	1022
Db	555	AGGACGGGGAGGAGCGGGAGGAGGACGAGGACGGGGAGGAGGAGAGAGAGAGAGACGGGGAGGACG	496
QY	1023	AGGAGAGAGAGCGCGTGAAGTGAATTTCCGCGGTGCCAAGACCGCGTGTGCTGCGCGCA	1082
Db	495	GGGAGGACGGGGAGGAGGACGAGGACGGGGAGGACGGGGAGGAGGAGGACGAGGACGGGGAGG	436
QY	1083	CGGGGCGGAACATGAGCATCCCGGCCATACGCTTCAAGATCAGGCGGACGCTGAGGAGA	1142
Db	435	AGGACGAGGACGGGGAGGACGGGGAGGACGGGGAGGAGGAGCAGAGGACGGGGAGGAGGACG	376

[illegible]

RESULT 3
US-09-197-649-7
; Sequence 7, Application US/09197649

```

1 GENERAL INFORMATION:
2 APPLICANT: Gold, Larry
3 APPLICANT: Tuerk, Craig
4 APPLICANT: Pribnow, David
5 APPLICANT: Smith, Jonathan D.
6 TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
7 FILE REFERENCE: NEXX02/C1-CON
8 CURRENT APPLICATION NUMBER: US/09/197,649
9 CURRENT FILING DATE: 1998-11-23
10 EARLIER APPLICATION NUMBER: 07/829,461
11 EARLIER FILING DATE: 1992-01-31
12 EARLIER APPLICATION NUMBER: 07/739,055
13 EARLIER FILING DATE: 1991-08-01
14 EARLIER APPLICATION NUMBER: 07/561,968
15 EARLIER FILING DATE: 1990-08-02
16 NUMBER OF SEQ ID NOS: 26
17 SOFTWARE: PatentIn Ver. 2.0
18 SEQ ID NO 7
19 LENGTH: 390
20 TYPE: DNA
21 ORGANISM: Artificial Sequence
22 FEATURES:
23 OTHER INFORMATION: Description of Artificial Sequence: Sequence
24 OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
25 OTHER INFORMATION: fragments having NcoI restriction sites.
26 US-09-197-649-7

```

Query Match	0.94;	Score 95.4;	DB 3;	Length 390;
Best Local Similarity	53.34;	Pred. No. 5.3e-12;		
Matches .201; Conservative	0;	Mismatches 176;	Indels 0;	Gaps 0;

[illegible]


```

1  APPLICANT: FRASER, Claire M.
2  APPLICANT: VENTER, John C.
3  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
4  TITLE OF INVENTION: TUBERCULOSIS
5  FILE REFERENCE: 24366-20007.00
6  CURRENT APPLICATION NUMBER: US/09/103,840A
7  CURRENT FILING DATE: 1998-06-24
8  NUMBER OF SEQ ID NOS: 2
9  SOFTWARE: PatentIn Ver. 2.1
10 SEQ ID NO 2
11 LENGTH: 4403765
12 TYPE: DNA
13 ORGANISM: Mycobacterium tuberculosis
14 FEATURE:
15 OTHER INFORMATION: CDC 1551
16 OTHER INFORMATION: "n" bases at various positions throughout the sequence
17 OTHER INFORMATION: represent a, t, c or g
18 US-09-103-840A-2

```

Query Match	0.84	Score 91.4	DB 3	Length 4403765
Best Local Similarity	46.64	Pred. No. 3.6e-08		
Matches 293; Conservative	0	Mismatches 336	Indels 0	Gaps 0

[illegible]

RESULT 8
US-09-902-540-1357
Sequence 1357, Application US/05902540
Patent No. 6833447
GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Miegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1357
LENGTH: 612
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(612)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 0.8%; Score 87.4; DB 4; Length 612;
Best Local Similarity 50.1%; Pred. No. 5.4e-10;
Matches 214; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

QY 4598 AACAAATTTTAAACAGCAATTTTGAATTTGACAAATTTTAAACGGGTAT 4657
DB 143 AAAT 202
QY 4658 TCCTGACATTTTCAAAATTTGATCAAAATTTTAAACGACTTTCTCAATTTGA 4717
DB 203 AATTA 262
QY 4718 GCATATTTTAAATTTAAAGTTCACAAATTTGCACTTTTAAATTTAGCGAGAA 4777
DB 263 AATAAAATTTAAATAATTTAAATTTTAAATAATTTAAATAATTTAAATTA 322
QY 4778 CATTTGAAATTTTAAATTTTGAATTTTGAATTTTCTATTTCTGACAAAT 4837
DB 323 AA 382
QY 4838 TGAATAATGACGATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 4897
DB 383 AATAAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 442
QY 4898 ACATATTTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4957
DB 443 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 502
QY 4958 ACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5017
DB 503 AA 562
QY 5018 GGAATAA 5024
DB 563 AAAAAAA 569

RESULT 9
US-09-056-556-182
Sequence 182, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yaelir A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington

TREATY

COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 182:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-056-556-182

Query Match 0.7%; Score 78.6; DB 3; Length 985;
Best Local Similarity 48.5%; Pred. No. 8.8e-08;
Matches 311; Conservative 0; Mismatches 319; Indels 11; Gaps 3;

QY 649 CATTGCCACCTGATCCGACGACGACGAGTCCGCGCAACGCTTGCCTGCGCAGC 708
DB 3 CAGCGCTACCGGTGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCG 62
QY 709 CGCGACGCGCAAGACGACGCGCATGCGGTGCGCGCGCGCGCGCGCGCGCG 768
DB 63 CAACAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122
QY 769 CCATGCGGTGATCCGACGCGGACGCGCAAGACGCGCATGCGGTGCGCGCG 828
DB 123 CGGACCGGAGGTCCGCGCGGTCCGCGCGCGCGCGCGCGCGCGCGCGCG 182
QY 829 GAACGCGCAAGACGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCG 888
DB 183 TGGCG 239
QY 889 GAACGCGCAAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948
DB 240 TACCG 299
QY 949 CGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1002
DB 300 CG 359
QY 1003 -GCATGCG 1061
DB 360 TGACG 419
QY 1062 AGGACGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1121
DB 420 CG 479
QY 1122 TCAGCGCGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1181
DB 480 AGGTCTCG 538
QY 1182 GGTCTGTTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1241
DB 539 ACAGCG 598
QY 1242 TGCGCACG 1282
DB 599 ATGCG 639

RESULT 10
US-09-072-596-177
Sequence 177, Application US/09072596
Patent No. 6458366
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mark, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 177:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-596-177

Query Match 0.7%; Score 78.6; DB 3; Length 985;
Best Local Similarity 48.5%; Pred. No. 8.8e-08;
Matches 311; Conservative 0; Mismatches 319; Indels 11; Gaps 3;

QY 649 CATGGCCACCGCTAGCCGAGAGCAAGAGTGGCCGCGCAACGGCTTGGCCGCGCGCAGC 708
DB 3 CAGGCGCTACCGGTGGCGCGCGGCTTCCGCGCGCGCGCGAGAGCGCGAGCGG 62
QY 709 CGCGAAGCGCAAGAGCAAGCGCCATGCGGTGCTGCGCGCGCTGAAACGGCAAGAGCAAGCG 768
DB 63 CAACAGCGGTGTGGCGCGCGCAACAGCGTCCGCGCGCGCGCGCGGTGCAAGCGCGCAAGGG 122
QY 769 CCATGGCGTGGATGCCGACGCGGAAACGGCAAGAGCAAGCGCCATGCGGTGCTGCGCAAGC 828
DB 123 CGGCAACCGGAGGTCCCGCGCGGTCCGCGCGAGCAACCCCGCTGCTGCTTCCGCGG 182
QY 829 GAACGGCAAGAGCAAGCGCCATGCGGAGCACTGCGCGCAACGGCGCGAGCGCACTGC 888
DB 183 TGCGCGCGCGCGCAAGG---TGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
QY 889 GAACGGCAAGAGCAAGCGCGCAACCGCGAGAGCAAGCGCGCATGAGCGCGCGCGCAAG 948
DB 240 TACCGCGCGCGCGCGCGCGCGCTTGTGCGCGCGCAACCGGTAGTCAAGCATCGCGCGCGCGCG 299

QY 949 CGCGAGAGCAAGCATCCGAGACTCCGCGCGCGAAGCGCGAGCAAGCG----- 1002
DB 300 CGCGCGCGCGGTGACCGCGCGCGCATGGGGCGACGCGTCTGGGCTTGGGCTTCCGCGCTT 359
QY 1003 -GCATGCGCGCGCGCGCGCGCAAGAGAGAGAGCGCGGTGAGTGTGAATTTCCGCGGTGCA 1061
DB 360 TGAACGGCGCGCAAGCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCATCA 419
QY 1062 AGGACGCGGTGCTGCGCGCGCAAGCGCGCGCGCAATGACATCCGCGCGCATACGATCAAGA 1121
DB 420 CGGGCGCGCGCGCGCGCGCGCGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479
QY 1122 TCAGCGGAGCGGTGACAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1181
DB 480 AGGTCTCGCGCAACAGCGCGCGGTGCGCGGTGACGCGGTGAGCGCGCGCGCGCGCGCG 538
QY 1182 CGTCCGTGTTCCCGGCTTCCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1241
DB 539 ACAGCGCGCAACGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598
QY 1242 TGCGCACCGGCGCAAGTCACTGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1282
DB 599 ATGCGCGCAACG 639

RESULT 11
US-09-072-967-182
Sequence 182, Application US/09072967
Patent No. 6592877
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mark, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 182:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-967-182

Query Match 0.7%; Score 78.6; DB 4; Length 985;
Best Local Similarity 48.5%; Pred. No. 8.8e-09;
Matches 311; Conservative 0; Mismatches 319; Indels 11; Gaps 3;

QY 649 CATGCGCAACCGTACCGCGAGAGCGAGTGCAGCGGAGCGGCTTGCCTGAGCGCGAGC 708
Db 3 CAGCGCTACCGGTGCGCGCGGTTCCCGCGCGCGCGCGAGAGCGAGCGCGG 62
QY 709 CGCGAAGCGCAAGCAAGCGGCTATGCGTGTGCGCGCGGCGGTAAGCAAGCAAGC 768
Db 63 CACAGCGGTGTGCGCGCGCAACCGGCTCCGCGCGCGCGCGGTGAGCGCGAGCGG 122
QY 769 CCATGCGGTGATGCGCGCGAGCGGCAAGCGGCAAGCGGCGGCGGCTGCGCGAGCGC 828
Db 123 CGGACCGGAGGTGCGCGCGGTCCGCGCGGTCCGCGCGGACCGCGCGGTGCGCGG 182
QY 829 GAACGCGAAGCAAGCGGCTATGCGGAGCGGCAAGCGGCGCGCGCGCGCGCGCG 888
Db 183 TGGCGCGCGCGCGCGAGG---TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
QY 889 GAACGCGAAGCAAGCGGCTATGCGGAGCGGCAAGCGGCGCGCGCGCGCGCGCG 948
Db 240 TACCGCGCGCGCGCGCGCGGTGTGCGCGCGCGCGGTGCGCGCGCGCGCGCGCG 299
QY 949 CGGCGAAGCAAGCGGCTATGCGGAGCGGCTCCGCGCGCGCGCGCGCGCGCGCG 1002
Db 300 CGGCGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
QY 1003 -GCATGCGCGCGCGCGCGCGCGAGAGCGGAGCGGCGGTGAGATTTGCGCGGTGCGCA 1061
Db 360 TGAAGCG 419
QY 1062 AGAAGCGCGGTGCG 1121
Db 420 CGGCG 479
QY 1122 TCAGCGCGAGGTGCGCGCGAGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1181
Db 480 AGGTCTCGCGCGCAACCGCGCGGTGCGCGCGGTGCGCGCGCGCGCGCGCGCGCG 538
QY 1182 CGTCTGCTGCTCCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1241
Db 539 ACG 598
QY 1242 TGCGCAAGCG 1282
Db 599 ATGCG 639

RESULT 12
US-09-489-847-67
Sequence 67, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489, 847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 67
LENGTH: 2434
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (27)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (73)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (75)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (130)
OTHER INFORMATION: n equals a,t,c,g, or c
US-09-489-847-67

Query Match 0.7%; Score 77; DB 4; Length 2434;
Best Local Similarity 54.4%; Pred. No. 4e-07;
Matches 178; Conservative 0; Mismatches 145; Indels 4; Gaps 1;

QY 4661 TGAACATTTTCAAAATTTGATCAAAATTTTAAAGACCTCTTCTCAATTTGAGCA 4720
Db 2110 TGATCATTTTACAGAAATTTATGAAAGATTTTGAATTTTGTGATGCT 2169
QY 4721 ATATTTAAATTTATTAATAAGTCAACATTTTGAATTTTAAATTTAGCGAGACAT 4780
Db 2170 ATGTTAGCGGAAACCAATTTTATGATTTTAAACCTTGATGAGAAACATTTGACA 2229
QY 4781 TTGAATTTCTAAATTTTGAATTTGAAACATTTTCTATTTCTGAACAAATTTGA 4840
Db 2230 TGTAATATGCTCAACTTTCTCAATTTTCTGTAATTTTCTAAGATACATTTAAATGTT 2289
QY 4841 AAATACAGCTAATTTGATTAATTTTGAATAATGCAATTTTGAATTTTGAACA 4900
Db 2290 TTATAT---TTTTTTTAAAGTAAATGACCGAGTAAAGAAATTAATTAACAGACA 2345
QY 4901 TATTTGAAAAAACAATACTTTAAAGTAAATTAATAATTAATAATTAATAATTAAGACA 4960
Db 2346 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2405
QY 4961 TAAAAATTAAGCAAAAAATAAAGAAA 4987
Db 2406 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2432

RESULT 13
US-09-902-540-1280
Sequence 1280, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

```

: FILE REFERENCE: 38-10115849JB
: CURRENT APPLICATION NUMBER: US/09/902,540
: CURRENT FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: 60/217,883
: PRIOR FILING DATE: 2000-07-10
: NUMBER OF SEQ ID NOS: 16825
: SEQ ID NO 1280
: LENGTH: 1039
: TYPE: DNA
: ORGANISM: Myxococcus xanthus
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(1039)
: OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

```

Query Match	0.7%	Score 75.6;	DB 4;	Length 1039;
Best Local Similarity	48.3%;	Pred. No. 4.6e-07;		
Matches 210; Conservative	0;	Mismatches 225;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 14
US-10-148-806-3/c
; Sequence 3, Application US/10148806
; Patent No. 6762042
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNAA
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/148, 806
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: 66/169, 970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ. ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3

```

```

; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-148-806-3

```

Query Match	0.7%	Score 75.4	DB 4	Length 114793
Best Local Similarity	48.2%	Pred. No. 1.5e-05		
Matches 274; Conservative	0	Mismatches 291	Indels 4	Gaps 22

QY	673	CGAAGTGCAGCGCAACGGCTTTCGCTTGCGCCGACGCGCGCAACGGCGAAGGCAACGGCCA	732
Db	31497	CGGCTGCGCGAAGAGGTGCGGGGGCGGGGCGGGGCGGAGCTCGAGCGGGCAGGGGCCG	31438
QY	733	TGGCGTGGCTGCCCGGTGAAACGGCAAGACCAACGGCCATGGCGTGGATGCCGACCGCA	792
Db	31437	GGGCGGGGCGCGGGCGAAGCTCGGACCGGGCAAGGGGCGGGGCGGGGCGGGGCGGAGCT	31378
QY	793	CGCGAAGACCAACGGCCATGGCGTGTGCTCGACCGCAACGGCAAGACCAACGGCCATGC	852
Db	31377	CGAGCGGGGCGAGGGGCGGGGCGGGGCGGGGCGGAGCTCGAGCGGGCGAGGGCGGGG	31318
QY	853	CGAGGGCACTGCGGACCGGCGAGGGCACTGGGAAACGGCAACCAACGGGCCACCG	912
Db	31317	CGGGGCGAGGGCTTAA--GCAAGCTAGAGGGGCGGGGCACTGGGCGGGGCGGGGCGAG	31267
QY	913	CGAAGCAACCGGCGCATGCTGAGGGCGGCGCAACGGCAACGGCGAAGCAACGAGCATGCGCA	972
Db	31260	GCGGAGCTCGGAGCGGGGCAAGGGGCGGGGCGGGGCGGGGCGGAGCTCGGAGCGGGCAAG	31207
QY	973	GGAATCCGCGCGCGCAACGGCGAAGACCAACGGGCACTGCGGCGGCGGCGGCGAAGGAGGA	1032
Db	31200	GGCGGGGGCGG--GGGCGAGGCTTAGCAGCAGGCGAGGGGGCGGGGCACTGGGCGGGGGCGG	31142
QY	1033	GGCGGTGAGGTGAATTTGCGGGGTGCCAAAGGACGGGCGTGTGTCGGGCGGCGACGGGGCGAA	1092
Db	31141	GGCCAGGGCGGAGCTCGAGAGGGGCGAGGGGCGGGGCGGGGCGAGGCTTCGAGAC	31087
QY	1093	CATGAGCATCCGGCGCATACGATACAGATCAAGCGGACGTTGCAAGAAAGGGCGGG	1152
Db	31081	GGGCGAGGGGCGGGGGGGGGGGGCAAGGCTTAGCAGACGAGGGGGCGGGGCACTGGGCGGG	31022
QY	1153	GGCGGTGCTGCGGCTGGCCCAAGGGGACCGGTCCGTTCCTGGGCTTCGCAACGGCGGT	1212
Db	31021	GGCGGGGCTTAGGGCGGAGCTCGGAGGGGCGCAAGAGCGGGCGGGGCACTCGAGGCGCTG	30962
QY	1213	CGAGGCGGAGAACGCGCTGCGCGGCGCGC	1241
Db	30961	AGCTGAGGTGAGGGTGGGGGGCTCTCTCC	30933

```

RESULT 15
US-09-806-708B-22/c
: Sequence 22, Application US/09806708B
: Patent No. 6784342
: GENERAL INFORMATION:
: APPLICANT: The University of British Columbia
: TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
: FILE REFERENCE: 4810-58741
: CURRENT APPLICATION NUMBER: US/09/806,708B
: CURRENT FILING DATE: 2001-04-03
: PRIOR APPLICATION NUMBER: US 60/147,133
: PRIOR FILING DATE: 1999-08-04
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 22
: LENGTH: 1141
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: NAME/KEY: promoter
: LOCATION: (1)..(1141)
: OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoter
US-09-806-708B-22

```


Db 15755 CTGGAGACAGTCGACGCCACA 15733

RESULT 4
US-09-544-3988-9/c

Sequence 9, Application US/095443988
Patent No. 6770461
GENERAL INFORMATION:
APPLICANT: Carulli, John P.
APPLICANT: Little, Randall D.
APPLICANT: Recker, Robert R.
APPLICANT: Johnson, Mark L.
TITLE OF INVENTION: High bone mass gene of 11q13.3
FILE REFERENCE: 032796-013
CURRENT APPLICATION NUMBER: US/09/544,3988
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/229,319
PRIOR FILING DATE: 1999-01-13
PRIOR APPLICATION NUMBER: US 60/071,449
PRIOR FILING DATE: 1998-01-13
PRIOR APPLICATION NUMBER: US 60/105,511
PRIOR FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 641
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 72049
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (8356), (8385), (38585)
OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-544-3988-9

Query Match 75.8%; Score 18.2; DB 4; Length 72049;
Best Local Similarity 87.0%; Pred. No. 56;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCGATCCACA 24

Db 19346 CTGGAGACAGTCGACGCCACA 19324

RESULT 5
US-09-543-771B-9/c

Sequence 9, Application US/09543771B
Patent No. 6780609
GENERAL INFORMATION:
APPLICANT: Carulli, John P.
APPLICANT: Little, Randall D.
APPLICANT: Recker, Robert R.
APPLICANT: Johnson, Mark L.
TITLE OF INVENTION: High bone mass gene of 11q13.3
FILE REFERENCE: 032796-014
CURRENT APPLICATION NUMBER: US/09/543,771B
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 09/229,319
PRIOR FILING DATE: 1999-01-13
PRIOR APPLICATION NUMBER: US 60/071,449
PRIOR FILING DATE: 1998-01-13
PRIOR APPLICATION NUMBER: US 60/105,511
PRIOR FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 641
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 72049
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (8356), (8385), (38585)
OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown

US-09-543-771B-9

Query Match 75.8%; Score 18.2; DB 4; Length 72049;
Best Local Similarity 87.0%; Pred. No. 56;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCGATCCACA 24
Db 19346 CTGGAGACAGTCGACGCCACA 19324

RESULT 6
US-09-949-016-14199/c

Sequence 14199, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14199
LENGTH: 140844
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(140844)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14199

Query Match 75.8%; Score 18.2; DB 4; Length 140844;
Best Local Similarity 87.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCGATCCACA 24

Db 87157 CTGGAGACAGTCGACGCCACA 87135

RESULT 7
US-08-422-699A-8/c

Sequence 8, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Houman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-699A-8

Query Match 74.2%; Score 17.8; DB 2; Length 2511;
Best Local Similarity 90.5%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGAGTTCAGTCGATCCACA 24
DB 2407 GGGGTCTCAGTCGATCCAAAA 2387

RESULT 8
US-08-422-706B-8/c
Sequence 8, Application US/08422706B
Patent No. 597733
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-706B-8

Query Match 74.2%; Score 17.8; DB 2; Length 2511;
Best Local Similarity 90.5%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGAGTTCAGTCGATCCACA 24
DB 2407 GGGGTCTCAGTCGATCCAAAA 2387

RESULT 9
US-08-422-699A-12/c
Sequence 12, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-422-699A-12

Query Match 74.2%; Score 17.8; DB 2; Length 2726;
Best Local Similarity 90.5%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTCGATCCAA 24
DB 2660 GGGGCTCAGTCGATCCAAA 2640

RESULT 10
US-08-422-706B-12/c
Sequence 12, Application US/08422706B
Patent No. 5977333
GENERAL INFORMATION:
APPLICANT: Broome, J. David
APPLICANT: Houseman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Hartley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-422-706B-12

Query Match 74.2%; Score 17.8; DB 2; Length 2726;
Best Local Similarity 90.5%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTCGATCCAA 24
DB 2660 GGGGCTCAGTCGATCCAAA 2640

RESULT 11
US-08-484-044-11/c
Sequence 11, Application US/08484044
Patent No. 5552282
GENERAL INFORMATION:
APPLICANT: Caskey, C. T.
APPLICANT: Fu, Ying-Hui
APPLICANT: Friedman, David L.
APPLICANT: Pizzuti, Antonio
APPLICANT: Fenwick, Raymond G.
TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pulbright & Jaworski, L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,044
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/019,940
FILING DATE: 19-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5443
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 3182 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-484-044-11

Query Match 74.2%; Score 17.8; DB 1; Length 3182;
Best Local Similarity 90.5%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTGCATCCAA 24
DB 3104 GGGGCTCAGTGCATCCAAA 3084

RESULT 12
US-08-422-699A-10/C
Sequence 10, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545

FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-9540
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 3323 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc difference
LOCATION: replace(518..3323, "")
OTHER INFORMATION: /standard_name="cDNA 41"
FEATURE:
NAME/KEY: misc difference
LOCATION: replace(769..3323, "")
OTHER INFORMATION: /standard_name="cDNA 28"
US-08-422-699A-10

Query Match 74.2%; Score 17.8; DB 2; Length 3323;
Best Local Similarity 90.5%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTGCATCCAA 24
DB 3219 GGGGCTCAGTGCATCCAAA 3199

RESULT 13
US-08-422-706B-10/C
Sequence 10, Application US/08422706B
Patent No. 5977333
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/264,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 3323 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(518..3323, "")
OTHER INFORMATION: /standard_name= "CDNA 41"
NAME/KEY: misc_difference
LOCATION: replace(769..3323, "")
OTHER INFORMATION: /standard_name= "CDNA 28"
US-08-422-7068-10

Query Match 74.2%; Score 17.8; DB 2; Length 3323;
Best Local Similarity 90.5%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGATCTCAGTCATCCACAA 24
Db 3219 GGGGTCTCAGTCATCCAAA 3199

RESULT 14
US-08-484-044-10/C
Sequence 10, Application US/08484044
Patent No. 5552282
GENERAL INFORMATION:
APPLICANT: Caskey, C. T.
APPLICANT: Fu, Ying-Hui
APPLICANT: Friedman, David L.
APPLICANT: Pizutti, Antonio
APPLICANT: Fenwick, Raymond G.
TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,044

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/019,940
FILING DATE: 19-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5443
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 11613 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-484-044-10

Query Match 74.2%; Score 17.8; DB 1; Length 11613;
Best Local Similarity 90.5%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGATCTCAGTCATCCACAA 24
Db 11107 GGGGTCTCAGTCATCCAAA 11087

RESULT 15
US-08-261-822A-9
Sequence 9, Application US/08261822A
Patent No. 5650553
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph R. et al.
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,822A
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Beardsell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1722 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-261-822A-9

Query Match 73.3%; Score 17.6; DB 1; Length 1722;
 Best Local Similarity 83.3%; Pred. No. 61;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCTGGGATCTCAATGCATCCACCA 24
 Db 1122 TTGGGATGCCAATGCATCCACCA 1145

Search completed: October 15, 2005, 03:25:12
 Job time : 5.70591 secs

```

FEATURES
    - source
        Contact: "Wing RA"
        Clemson University Genomics Institute
        Clemson University
        100 Jordan Hall, Clemson, SC 29634, USA
        Tel: 864 656 7288
        Fax: 864 656 4293
        Email: rwing@clemson.edu
        Total hg bases = 402
        Seq primer: AATTACCTCCTACTAAGG
        High quality sequence stop: 597.
        Location/Qualifiers
            1. 681
                /organism="Hordeum vulgare subsp. vulgare"
                /mol_type="mRNA"
                /cultivar="Morex"
                /sub_species="vulgare"
                /db_xref="taxon:112509"
                /clone="HVSME0025M09f"
                /rname_type="green seedling leaf"
                /lab_host="TTC121"
                /clone_1b="Hordeum vulgare green seedling EST library
                HVCNDN0014 (Blumeria infected)"
                /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                XhoI; Morex (mla) plants were greenhouse grown in the R
                Wise lab at Iowa State University, Ames, IA; 7 day old
    
```



```

QY 1717 CGACATCGACTGCTGATCCATCGCGCAAGAACACCCAGCCGATGTCATATA 1776
DB 290 CGACATCGACTGCTGATCCATCGCGCAAGAACACCCAGCCGATGTCATATA 349
QY 1777 CCCCAACAACCCGTCGGCGAGGCTTTACTCTCTACGACCATCTGTCCAGGTTTCATCC 1836
DB 350 CCCCAACAACCCGTCGGCGAGGCTTTACTCTCTACGACCATCTGTCCAGGTTTCATCC 396
QY 1837 TTGCTGCTGTAATATGATTCAGTTCAGTGCACCTGTCGTAATCTTTTGGCAATCC 1896
DB 397 -----
QY 1897 ATACTGACTGATGTTGCTCAATTAGTCCGCGAGGTGCGAAAGGCTCGAATATGCT 1956
DB 397 -----AGTCCGCGAGGTGCGAAAGGCTCGAATATGCT 433
QY 1957 GATTGCTGACGAGTATACGCGAAGCTGTTCTGCGGACGCGCCGTTTCATCCATGCG 2016
DB 434 GATTGCTGACGAGTATACGCGAAGCTGTTCTGCGGACGCGCCGTTTCATCCATGCG 493
QY 2017 AGTGTGTCGACATCACCCGTCGTCCTCCATAGGCTCTGTCGCAACTCATGATTA 2076
DB 494 AGTGTGTCGACATCACCCGTCGTCCTCCATAGGCTCTGTCGCAACTCATGATTA 553
QY 2077 GCGTGGATGCGCGCTTGATGAGTGGTAGCGGTGTACGACCCGAGAAAGATCTTACAGGAAC 2136
DB 554 GCGTGGATGCGCGCTTGATGAGTGGTAGCGGTGTGTACGACCCGAGAAAGATCTTACAGGAAC 613
QY 2137 TAAAGTACTTAATC 2151
DB 614 TAAAGTACTTACATC 628

```

```

RESULT 3
LOCUS BF268125 770 bp mRNA linear EST 23-OCT-2001
DEFINITION HV_CEA0019P10F Hordeum vulgare seedling green leaf EST library
            HVCNA0004 (Blumeria challenged) Hordeum vulgare subsp. vulgare
            cDNA clone HV_CEA0019P10F, mRNA sequence.
ACCESSION BF268125
VERSION BF268125.2 GI:13263792
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
          Eurycotla: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
          Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 770)
          Wing, R., Close, T. J., Kleinof, A., Wise, R., Wei, F., Begum, D.,
          Fritsch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J.,
          Choi, D. W., Fenton, R. D., Oates, R. and Main, D.
          Development of a genetically and physically anchored EST resource
          for barley genomes: Blumeria infected incompatible (Mla13)
          unpublished (2001)
          Unpublished (2001)
JOURNAL On Nov 17, 2000 this sequence version replaced gi:11199120.
COMMENT Contact: Wing RA
          Clemson University Genomics Institute
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Total hg bases = 586
          Seq primer: AATTAACTCTCACTAAAGG
          High quality sequence stop: 611.
          Location/Qualifiers
            1..770
              /organism="Hordeum vulgare subsp. vulgare"
              /mol_type="mRNA"
              /cultivar="C16155 (Mla13)"
              /sub_species="vulgare"
              /db_xref="taxon:112509"

```

ORIGIN

```

Query Match 4.0%; Score 439.6; DB 2; Length 770;
Best Local Similarity 83.6%; Pred. No. 2e-78;
Matches 565; Conservative 0; Mismatches 14; Indels 97; Gaps 2;

```

/clone="HV_CEA0019P10F"
 /issue_type="seedling green leaf"
 /lab_host="TUC121"
 /clone_1lib="Hordeum vulgare seedling green leaf EST
 library HVCNA0004 (Blumeria challenged)"
 /note="Vector: lambda2AP; Site 1: BcORI; Site 2: XhoI;
 C.I. 16155 (Mla13) plants are greenhouse grown in the R
 Wise lab at Iowa State University, Ames, IA; 7 day old
 green seedlings were challenged with isolate A27
 (AVZM1a13) of Blumeria graminis f. sp. hordei, and leaves
 were harvested 20 and 24 hr post-inoculation and snap
 frozen; uninoculated leaves were harvested 20 hr
 post-inoculation (Wei, Wise). In the TJ close lab at the
 University of California, Riverside, total RNA was
 prepared from each sample pool, equal quantities of all
 three RNA pools were combined, poly(A) RNA was purified
 from the mixture, one cDNA library was made, and 1 million
 pfu were in vivo excised to give pluscript SK(-) cDNA
 phagemids (Choi, Close). Phagemids were plated and picked
 at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Fritsch, Atkins and Wing). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Wing, Yu, Fritsch, Henry, Simmons,
 Oates, Rambo, Main). The sequence has been trimmed to
 remove vector sequence and contains a minimum of 100 bases
 of phred value 20 or above. For more details on library
 preparation and sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
 this clone see <http://www.genome.clemson.edu/orders> Also
 see Close TJ, Wing R, Kleinof A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomes. Barley Genetics Newsletter 31:29-30.
 (<http://wheat.pw.usda.gov/g99pages/bgn/31/cover.html>)"

```

QY 1477 ACAAGCGCGCGTGGAGACACCTTCCGAGGCGCTCCGTCATGTCATGCGCGAGCA 1536
DB 101 ACAAGCGCGCGTGGAGACACCTTCCGAGGCGCTCCGTCATGTCATGCGCGAGCA 160
QY 1537 GCTTCTTCAACCGCGCGGAGACCCAGGCGATGAGGTCATTAATCCGCGTCTGAGCCA 1596
DB 161 GCTTCTTCAACCGCGCGGAGACCCAGGCGATGAGGTCATTAATCCGCGTCTGAGCCA 220
QY 1597 GACCGCGCGCGCAACATTTGCTCCCGAGGCGATACCCAACTACGAGCGCGCGC 1656
DB 221 GACCGCGCGCGCAACATTTGCTCCCGAGGCGATACCCAACTACGAGCGCGCGC 280
QY 1657 GCGCTTCAACAGGCTGAGGTCGCGCATTTGACTTCCCGACAGAGGGTGGAGAT 1716
DB 281 GCGCTTCAACAGGCTGAGGTCGCGCATTTGACTTCCCGACAGAGGGTGGAGAT 340
QY 1717 CGACATCGACTGCTGATCCATCGCGCAAGAACACCCAGCCGATGTCATATA 1776
DB 341 CGACATCGACTGCTGATCCATCGCGCAAGAACACCCAGCCGATGTCATATA 400
QY 1777 CCCCAACAACCCGTCGGCGAGGCTTTACTCTCTACGACCATCTGTCCAGGTTTCATCC 1836
DB 401 CCCCAACAACCCGTCGGCGAGGCTTTACTCTCTACGACCATCTGTCCAGGTTTCATCC 447
QY 1837 TTGCTGCTGTAATATGATTCAGTTCAGTGCACCTGTCGTAATCTTTTGGCAATCC 1896
DB 448 -----
QY 1897 ATACTGACTGATGTTGCTCAATTAGTCCGCGAGGTGCGAAAGGCTCGAATATGCT 1956
DB 448 -----AGTCCGCGAGGTGCGAAAGGCTCGAATATGCT 484
QY 1957 GATTGCTGACGAGTATACGCGAAGCTGTTCTGCGGACGCGCCGTTTCATCCATGCG 2016

```

Db 485 GATTGCTGACGAGGTATACGGCAAGCTGTTCTGGGACGCGCCCGTTATCCCAATGGG 544
Qy 2017 AGTGTGGGACATACCCCTGCTGCTGCATAGAGGTCTGTGCTCAAGTCAATGATAGT 2076
Db 545 AGTGTGGGACATACCCCTGCTGCTGCATAGAGGTCTGTGCTCAAGTCAATGATAGT 604
Qy 2077 GCGTGGATGGCGGCTTGATGGGTAGCGGTGACG-ACCCCAAGAAAGTCTTACAGGAAA 2135
Db 605 GCTCGAGGCGGCGCTTGATGGGTAGCGGTGACGAGAACCCCAAGAAAGTCTTCCGGGAAA 664
Qy 2136 CTAAAGCTACTTAATC 2151
Db 665 CTAAAGACTTTACATC 680

RESULT 4
LOCUS CD870152 692 bp mRNA linear EST 11-JUL-2003
DEFINITION AZ02.113J24F001128 AZ02 Triticum aestivum cDNA clone AZ02113J24,
mRNA sequence.
ACCESSION CD870152
VERSION CD870152.1 GI:32553968
KEYWORDS EST
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Triticum.
REFERENCE 1 (bases 1 to 692)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
FEATURES
SOURCE location/Qualifiers
1..692
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="rectal"
/db_xref="taxon:4565"
/clone="AZ02113J24"
/tissue_type="root"
/clone_lib="AZ02"

ORIGIN
Query Match 3.5%; Score 384.2; DB 6; Length 692;
Best Local Similarity 76.4%; Pred. No. 3.9e-67;
Matches 548; Conservative 0; Mismatches 38; Indels 111; Gaps 1;
Qy 6657 GCAGAGCGGATCTGTCGACGACGCGGCGGAGAAACAGATCCGGCGATACGATCA 6716
Db 1 GCAAGAGTGGCATCTGCGGACGAGACGGGTGCCAAGAACAGATCCGGCGATACGATCA 60
Qy 6717 AGATCAGGCGAGCGTGGAGAGAGCGGCGCGCGCTGCTGCTGCGCCACCGTG 6776
Db 61 AGATCAGGCGAGCGTGGAGAGAGCGGCGCGCGCTGCTGCTGCGCCACCGTG 120
Qy 6777 ACCGTCGCTGTTCCCGGCTTCCGACGCGCGTGCAGGCGGAGACCGCTGCGCGG 6836
Db 121 ACCGTCGCTGTTCCCGGCTTCCGACGCGCGTGCAGGCGGAGACCGCTGCGCGG 180
Qy 6837 CGCTGCGACCGCGCGATTCATCTGTAAGCGCGCGCGCTGCGCTTCCCGCCGACGAA 6896
Db 181 CGCTGCGACCGCGCGATTCATCTGTAAGCGCGCGCGCTGCGCTTCCCGCCGACGAA 239
Qy 6897 GGTAAACATTACAGCTTACCGGTATGTATGCGTGAAGATGATCGCGCGGTTACTTAC 6956

Db 240 ----- 239
Qy 6957 GTGCCGCGCGTGTCTTCCCGGTGGGTTCAAAATTTTAACTTTATAGTACTTAT 7016
Db 240 ----- 239
Qy 7017 AAAAACAACAGCGCCGTAGACAGCACTTGTCAAGGGCGTGCCTTACAGTATCGGC 7076
Db 240 -----AGCCTGTAGACAGCACTTGTCAAGGGCGTGCCTTACAGTATCGGC 289
Qy 7077 CGACGAGTCTTCTCAACCGCGCGGGAACACAGGCGATGAGTCAATATCCCGGTCT 7136
Db 290 CGACGAGTCTTCTCAACCGCGGTGAAACCAAGCAATGAGTCAATATCCCGGTCT 349
Qy 7137 GAGCCAGACTGCGCGCCCAACATATCTGTTCCCGGCGAGGCTATCCAAATTAAGAGGC 7196
Db 350 AGCCCAAACTGTCCGCGCCCAACATATCTGTTCCCGGCGAGGCTATCCGAATTAAGAGGC 409
Qy 7197 GCGAGCGGCTTCAACAGCTGAGGTCCGGCACTTGACCTCATCCCGACAAAGGGGTG 7256
Db 410 GCGAGCGGCTTCAACAGCTGAGGTCCGGCACTTGACCTTATCCCGGAGAGGGGTG 469
Qy 7257 GGAGTGCAGTGCATCTGCTGGAATCCATTCGCGACAAACACACCGGAGTGTAT 7316
Db 470 GGAGTGCAGTGCATCTGCTGGAATCCATTCGGAACAAACACCGGAGTGTAT 529
Qy 7317 CATTAACCAACAAATCCGTGCGGACGTTACTCTTACGACCATCTGCGCAAGT 7373
Db 530 TATTAACCAACAAATCCGTGCGGCACTTACTCTTACGACCATCTTGGCCAAAGT 586

RESULT 5
LOCUS CD864826 730 bp mRNA linear EST 11-JUL-2003
DEFINITION AZ02.001M03F000711 AZ02 Triticum aestivum cDNA clone AZ02001M03,
mRNA sequence.
ACCESSION CD864826
VERSION CD864826.1 GI:32548642
KEYWORDS EST
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Triticum.
REFERENCE 1 (bases 1 to 730)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
FEATURES
SOURCE location/Qualifiers
1..730
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="rectal"
/db_xref="taxon:4565"
/clone="AZ02001M03"
/tissue_type="root"
/clone_lib="AZ02"

ORIGIN
Query Match 3.5%; Score 383; DB 6; Length 730;
Best Local Similarity 69.7%; Pred. No. 6.9e-67;
Matches 674; Conservative 0; Mismatches 55; Indels 218; Gaps 2;
Qy 6682 GGGGCGAAGAACGATCCGGCGGATACGTTCAAAATATGACGCGCGCGTGGAGGAG 6741

QY 7622 GTCTGATAGTCTGATGCGACTTGATGGTGGCGGTGTACGACCCCAAAAGT 7681
 DB 521 GTCTGATAGTCTGATGCGACTTGATGGTGGCGGTGTACGACCCCAAAAGT 580
 QY 7682 TTAAAGAAAC 7693
 DB 581 TTAAAGAAAC 592
 RESULT 7
 LOCUS CC692577 980 bp DNA linear GSS 19-JUN-2003
 DEFINITION OGJ939TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0447G06,
 genomic survey sequence.
 ACCESSION CC692577
 VERSION CC692577.1 GI:32097353
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 980)
 Whitelew,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSes: OGJ939TV
 CONTACT: Cathy Whitelew
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelew@tigr.org
 Seq primer: TP
 Class: sheared ends.
 Location/Qualifiers
 source 1..980
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0447G06"
 /clone_1ib="ZM 0.7 1.5 KB"
 /note="Vector: pBSGK-1; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"
 ORIGIN
 Query Match 3.3%; Score 363.8; DB 9; Length 980;
 Best Local Similarity 69.1%; Pred. No. 5; 9e-63;
 Matches 577; Conservative 0; Mismatches 237; Indels 21; Gaps 5;

QY 7315 ATCAAAACCAACAAATCCGTGGGAGCGTTTACTCTACGACATCTGAGGTT 7374
 DB 463 ATCAAAACCAACCAACCTTGGGAGGTCTACACCCAGACATTTGGCCAGCA 522
 QY 7375 TTGATCCATGATCTCTGCTGCTGTGATGACCCGCTGTGTGAACATGATATGGA 7434
 DB 523 CGGCTAGCCTCTT-----ATCTTTCGCTGGCCATTTGGATTTGATGGATCGG 571
 QY 7435 TTGGTTTGTAAATCGTGTGCTGATGATGCTGTTTGGTTATC-AGTGGCGAGGTGCA 7493
 DB 572 ATCGTTTGTGTCTTAAAGCACTAGCTCTGATGATCTAGTGTGCGAGGTGCA 631
 QY 7494 AGGAAGCTCGAATATTTGATGATGCTGACGAGGTTTACGCAACTGTCTGCGGACG 7553
 DB 632 AGGAAGCTTGAATATGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 691
 QY 7554 GCCCGTTTATCCGATGGGCGCTTTTGGGCAATTGCCCCGCTTGTCCATTTGATCT 7613
 DB 692 ACCCTTTCGCTCCCATGGGTGTCTTGGCCACATTCGCCCTGTGTGACCATAGATCA 751
 QY 7614 CTGTCAAGTGTGATATGATGCTGATGAGGACCTTGATGGGTGGCGGTGATGACCC 7673
 DB 752 CTGTCAAGATGATATGATGCTGATGAGGACCTTGATGGGTGGCGGTGATGATGATCC 811
 QY 7674 ACAAGATTTTGAAGAAATCAAGTATGATGATGATGATGATGATGATGATGATGATG 7733
 DB 812 AACAGGTTCGAAGAAACCAAGTA--ATATATCATTAATGATCATATCTTATGCTCTT 868
 QY 7734 CTGTGGGATTTGATTTTGTCTTAAATTTGATGCTCTTGTATTCATATCTTACGT 7793
 DB 869 GTTTCAGTATGATGTTGCTAAGCT--GTTTCATCATGATTTTGTGAGATATTTGAT 925
 QY 7794 CTATTACGAATTTACTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 7848
 DB 926 GATACCAAACTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 980

RESULT 8
 LOCUS BQ466157 606 bp mRNA linear EST 30-MAY-2002
 DEFINITION HT01121T HT Hordeum vulgare subsp. vulgare cDNA clone HT01121
 5-PRIME, mRNA sequence.
 ACCESSION BQ466157
 VERSION BQ466157.1 GI:21273939
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Hordeum.
 1 (bases 1 to 606)
 Zhang,H., Pocotkina,E., Michalek,W., Meschke,W., Stein,N. and
 Graner,A.
 Barley ESTs from germinating seeds
 Unpublished (2002)
 CONTACT: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 606 Std Error: 0.00
 Plate: 1 row: 1 column: 21
 Seq primer: T3.
 Location/Qualifiers
 source 1..606
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultiyar="barke"
 /sub_species="vulgare"
 /db_xref="taxon:112509"

Db	225	GAGCAACGGGCACGGCCGAGCCGCTGCCCGCAACGGCCGAGCAACGGCCATGAGGTGGC	284
Oy	1015	GACGCGACAGC-----AGAGAGAGCGCGGTGAGTGAATTTCCGCGGTGCCAAGACGG	1066
Db	285	GCGCGCACTGGCGCGCGCCGAGAGAGCGCGTGGAGTGAATTTCCGCGCGGCGCAAGCGCGG	344
Oy	1069	CGTCGTGCGCGGAGACGGGGGCGCAATATAGCATCCGGGCGGANAACGGTCAAGATCAAGCGC	1122
Db	345	CGTCCTGCGCGGAGACGGGGGCGCAATATAGCATCCGGGCGGATACGGTACAAATATACGGC	404
Oy	1129	GAGCGTGAAGAGAGAGGGGCGCGCGCGCGTGTCTGCCTGCGTCCACCGGAGACCCGTCGGT	1188
Db	405	GAGCGTGAAGAGAGACGGGGCGCGCGCGCGTGTCTGCCTGCGCGCAACGGGACCCGTCGGT	464
Oy	1189	GTTCCCGGCGCTTCCGCAACGGCCGTGAGGCGCAAGACGCGCGTGCCTGCGCGCGCTGCGCAC	1248
Db	465	TTTCCCGGCGCTTCCGCAACGGCCGTGAGGCGCAAGACGCGCGTGCCTGCGCGCGCTGCGCAC	524
Oy	1249	CGGCGAGTTCAACTGCTAACCCCGCGGGGCTCGGCGCTCCCGCGGACGAAAG	1299
Db	525	GGGGAGTTCAACTGCTAACCCCGCGGGGCTCGGCGCTCCCGCGGACGAAAG	575

RESULT 10	CA002831	CA002831	354 bp	mRNA	linear	EST 23-OCT-2001
LOCUS	CA002831					
DEFINITION	HS08K11r HS <i>Hordeum vulgare</i> subsp. <i>vulgare</i> cDNA clone HS08K11					
	5-PRIME, mRNA sequence.					

ACCESSION	CA002831
VERSION	CA002831.1
KEYWORDS	GI:24279813 EST.

SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT
Hordeum vulgare subsp. vulgare	Hordeum vulgare subsp. vulgare	1 (bases 1 to 354)	Barley ESTs from germinating seeds	Unpublished (2002)	Contact: Stein Nils
Hordeum vulgare subsp. vulgare	Hordeum vulgare subsp. vulgare	Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner, A.			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Triticeae; Hordeum.					

FEATURES
Source Location/Qualifiers
 1. .354

```

/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:258040"
/db_xref="taxon:112509"
/clone="HS08K11"
/tissue_type="embryo + scutellum"
/dev_stage="0-16 hours after imbibition"
/lab_host="X110-Gold"
/clone_1fb="HS"
/note="vector: plasmid SK+, Site_1: EcoRI (5' end of cDNA), Site_2: XhoI (3' end of cDNA). Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."

```

ORIGIN	
Query Match	3.1%; Score 345; DB 6; Length 354;
Best Local Similarity	100.0%; Pred. No. 3.3e-59;
Matches 345; Conservative	0; Mismatches 0; Indels 0; Gaps 0.

OY	817	G G C T G C C G A C G C G A A C G G C A A G A G C A A C G G C C A T G C C G A G G C C A T G C G A A C G G C C A C G G	876
D b	10	G G C T G C C G A C G C G A A C G G C A A G A G C A A C G G C C A T G C C G A G G C C A T G C G A A C G G C C A C G G	69
OY	877	C G A G G C C A C T G C G A A C G G C A A G A C C A A C G G C C A C C G G C A G A G C A A C G G C C A T G C T G A G G C	936
D b	70	C G A G G C C A C T G C G A A C G G C A A G A C C A A C G G C C A C C G G C A G A G C A A C G G C C A T G C T G A G G C	129
OY	937	C G C G G A C G C G A A C G G C C A G A G C A A C G A G C A T G C C G A G A C T C C G C G G C G A A C G G C G A G	996
D b	130	C G C G G A C G C G A A C G G C C A G A G C A A C G A G C A T G C C G A G A C T C C G C G G A A C G G C G A G	189
OY	997	C A A C G G G C A T G C G G C G G C G G C G A G A G A G A G A G G C G G T G A G T G A A T T T C G C G G	1056
D b	190	C A A C G G G C A T G C G G C G G C G G C G A G A G A G A G A G G C G G T G A G T G A A T T T C G C G G	249
OY	1057	T G C C A A G A C G G C C T G C T G G C G G C G A C G G G G C G A A C A T A G C A T C C G G G C G A T P C G G T A	1116
D b	250	T G C C A A G A C G G C C T G C T G G C G G C G A C G G G G C G A A C A T A G C A T C C G G G C G A T A C G G T A	309
OY	1117	C A A G A T A G C C G A G C G T G C A G A G A A G G G G C C G C G G C C C T G C T	1161
D b	310	C A A G A T A G C C G A G C G T G C A G A G A A G G G G C C C G G C C C C T G C T	354

RESULT	11
AJ433473	
LOCUS	
DEFINITION	AJ433473 S00011 Hordeum vulgare cDNA clone S000110036B03fl, mRNA
ACCESSION	AJ433473
VERSION	AJ433473
KEYWORDS	EST.
SOURCE	Hordeum vulgare
ORGANISM	Hordeum vulgare

REFERENCE
TITLE
AUTHORS
JOURNAL
COMMENT

Unpublished (2002)
Contact: Schulman AH

Barley EST's
Saren, A.-M., Tanskanen, J., Paulin, L. and Schulman, A.H.
1 (bases 1 to 373)
Pooidae; Triticeae; Hordeum.
Eumarysta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.

FEATURES
Source Location/Qualifiers
1. .373

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query	Local Match	Similarity	Score	DB	Length
Db	Best Local Match	3.1%;	340.4;	1	373;
	Matches	341;	Pred. No. 2.9e-58;		
		Conservative	0;</		

```

Db      88  CACCCGCGGAGACCGACGCGATCGAGTCATTAATCCCGGTGCTGGCCCAAGACGCGCG 147
Qy      1606  CGCCCAACATTGCTGCTCCCAAGGCGAGGCTACCCAACTAGAGGCGCGCGCGCTTCAA 1665
Db      148  CGCCCAACATTGCTGCTCCCAAGGCGAGGCTACCCAACTAGAGGCGCGCGCGCTTCAA 207
Qy      1666  CAGGCTGAGAGTCGCGGACATTTGCACTCTCATCCCGGCAAGGGGTGGAGATGACATCGA 1725
Db      208  CAGGCTGAGAGTCGCGGACATTTGCACTCTCATCCCGGCAAGGGGTGGAGATGACATCGA 267
Qy      1726  CTGCTGGAATCATTCGCGGACAGAACACACCGGCATGTCATCATTAACCCCAACA 1785
Db      268  CTGCTGGAATCATTCGCGGACAGAACACACCGGCATGTCATCATTAACCCCAACA 327
Qy      1786  CCGGTCGCGGACGCTTACTCTTAGACCATCTGTCCAAAGT 1827
Db      328  CCGGTCGCGGACGCTTACTCTTAGACCATCTGTCCAAAGT 369

RESULT 12
LOCUS   CD878050 635 bp mRNA linear EST 11-JUL-2003
DEFINITION AZ04.101M17P011002 AZ04 Triticum aestivum cDNA clone AZ04101M17, mRNA sequence.
ACCESSION CD878050
VERSION   CD878050.1 GI:32561866
KEYWORDS  EST.
SOURCE    Triticum aestivum (bread wheat)
ORGANISM  Triticum aestivum
REFERENCE 1 (bases 1 to 635)
AUTHORS   Genoplante.
TITLE      Genoplante, a major partnership french program in plant genomes
JOURNAL   Unpublished (2003)
COMMENT    Contact: Genoplante
           Genoplante
           93, rue Henri Rochefort 91025 EVRY CEDEX France
           Tel: 33 1 69 47 54 00
           Fax: 33 1 69 47 54 10
           This sequence has been generated in the framework of the french
           plant genome programme 'Genoplante' (http://www.genoplante.com
           and http://genoplante-info.infobiolegen.fr).
FEATURES
  source          location/Qualifiers
  1..635
  /organism="Triticum aestivum"
  /mol_type="mRNA"
  /culтив="rectal"
  /db_xref="taxon:4565"
  /clone="AZ04101M17"
  /tissue_type="root"
  /clone_11b="AZ04"

ORIGIN
Query Match 3.1%; Score 340; DB 6; Length 635;
Best Local Similarity 90.8%; Pred. No. 3.8e-58;
Matches 374; Conservative 0; Mismatches 35; Indels 3; Gaps 1;

Qy      891  ACGGCAAGACCAACGCGACCGCGAGACCAACGCGCATGTCAGGCGCGCGCGAGAG 950
Db      96   ACGGCGCGCGCGCGCGCAACGCGCAAGACCAACGCGCGCATGTCAGGCGCGAGAG 155
Qy      951  GCGAGAGCAAGAGATGCTCGAGAGCTCGCGCGCAACGCGGAGCAACGCGGCTGCGG 1010
Db      156  GCGAGAGCAAGAGCGCGCGCGCGCTCTCTGTAAAGCGCGAGAGCAACGCGGCTGCGG 215
Qy      1011  CGGCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1067
Db      216  CGGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 275
Qy      1068  GCGGTGTCGCGCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1127

```

```

Db      276  GGTCTGCGCGCGAGACGCGGCGGAGACATACGATCCGGCGATACGTAACAAGTACGG 335
Qy      1128  CGAGGTGAGAGAGAGGCGCGCGCGCTGCTGCTGCGCGCAACGCGGAGACCGGTCCG 1187
Db      336  CGAGGTGAGAGAGAGGCGCGCGCGCTGCTGCTGCGCGCAACGCGGAGACCGGTCCG 395
Qy      1188  TGTTCGCGCGCTTCCGCAACGCGCGCTGAGAGGCGGAGAGGCGGCGCGCGCGCGGCA 1247
Db      396  TGTTCGCGCGCTTCCGCAACGCGCGCTGAGAGGCGGAGAGGCGGCGCGCGCGCGGCA 455
Qy      1248  CCGGCAAGTTCAGTCTACCCCGCGCGCTGCGGCTCCCGCGGAGAG 1299
Db      456  CCGGCAAGTTCAGTCTACCCCGCGCGCTGCGGCTCCCGCGGAGAG 507

RESULT 13
LOCUS   AJ485409 360 bp mRNA linear EST 24-MAY-2002
DEFINITION AJ485409 S00011 Hordeum vulgare cDNA clone S0001100055B10P1, mRNA sequence.
ACCESSION AJ485409
VERSION   AJ485409.1 GI:21201364
KEYWORDS  EST.
SOURCE    Hordeum vulgare
ORGANISM  Hordeum vulgare
REFERENCE 1 (bases 1 to 360)
AUTHORS   Saren, A.-M., Tanekannen, J., Paulin, L. and Schulman, A.H.
TITLE      Unpublished (2002)
JOURNAL   Contact: Schulman AH
COMMENT    Institute of Biotechnology
           University of Helsinki
           P.O. Box 56 (Valkankari 6A), University of Helsinki FIN-00014,
           Finland.
FEATURES
  source          location/Qualifiers
  1..360
  /organism="Hordeum vulgare"
  /mol_type="mRNA"
  /db_xref="taxon:4513"
  /clone="S0001100055B10P1"
  /dev_stage="Developing seed"
  /clone_11b="S00011"
  /note="12,15,18 days after pollination"

ORIGIN
Query Match 3.1%; Score 337; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.4e-57;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1491  CAGAGCACCTGTCGAGAGGCGTGCCTACAGCTATCGGCGGACAGACGCTTCTCACCG 1550
Db      1   CAGAGCACCTGTCGAGAGGCGTGCCTACAGCTATCGGCGGACAGACGCTTCTCACCG 60
Qy      1551  CCGGCGGACCCAGGCGATGAGAGTCATTAATCCCGTCTGCGCCAGACCGCGGCGCA 1610
Db      61   CCGGCGGACCCAGGCGATGAGAGTCATTAATCCCGTCTGCGCCAGACCGCGGCGCA 120
Qy      1611  ACATTGCTCTCCCAAGGCGAGGCTACCAACTAGAGGCGCGCGCGCTTCAACAGG 1670
Db      121  ACATTGCTCTCCCAAGGCGAGGCTACCAACTAGAGGCGCGCGCGCTTCAACAGG 180
Qy      1671  TGAAGTCCGAGATTTCAGCTCATCCCGCAAGAGGCTGGAGATCGACATCGACTGCG 1730
Db      181  TGAAGTCCGAGATTTCAGCTCATCCCGCAAGAGGCTGGAGATCGACATCGACTGCG 240
Qy      1731  TGAATCATTCGCGCAAGAAACACACCGCATGCTATATAAACCACCAACCCGCT 1790
Db      241  TGAATCATTCGCGCAAGAAACACACCGCATGCTATATAAACCACCAACCCGCT 300

```

```

Oy 1791 GCGGACGGTTTACTCTCAGACCATCTGTCCAAAGT 1827
Db 301 GCGGACGGTTTACTCTCAGACCATCTGTCCAAAGT 337

RESULT 14
CG033764/c 940 bp DNA linear GSS 19-AUG-2003
LOCUS PUFYB66TB.ZM.0.6.1.0.Zea.mays.genomic.clone.ZMWBTA0759K12,
DEFINITION genomic survey sequence.
ACCESSION CG033764
VERSION CG033764.1 GI:33905920
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 940)
AUTHORS White, J., Quackenbush, J., Van Aken, S., Uterback, T.,
Reinick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Benner, J.
Maize Genomics Consortium
Unpublished (2003)
JOURNAL Other_GSSs: PUFYB66TD
COMMENT Contact: Cathy White, cwhite@maizegenomics.org
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white@maizegenomics.org
Seq primer: TR
Class: sheared ends.
FEATURES
source location/Qualifiers
1..940
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMWBTA0759K12"
/clone_1lb="ZM.0.6.1.0.KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COR selected genomic DNA library"
ORIGIN
Query Match 3.0%; Score 325.6; DB 9; Length 940;
Best Local Similarity 70.0%; Pred. No. 3.6e-55;
Matches 483; Conservative 0; Mismatches 199; Indels 8; Gaps 3;
Oy 7012 CTTATTAACCAACAGCGCGGTAGACAGCATTTGTCAAGGCGTCCCTTACAAAGTA 7071
Db 804 CTTTATTAACCAACAGCGGTAGACAGCATTTGTCAAGGCGTCCCTTACAAAGTA 745
Oy 7072 TCGGCGGACGAGCTTCTCTCAACCGCGCGGAAATCAGGCGATGATATATCCCG 7131
Db 744 TCAACTGATGACATCTTCTCAACCGGTAGACAGCATTTGTGATGCTCA 685
Oy 7132 GTGTGCGGACGAGCTTCTCTCAACCGCGCGGAAATCAGGCGATGATATATCCCG 7191
Db 684 GTCTTGTGCGGACGAGCTTCTCTCAACCGGTAGACAGCATTTGTGATGCTCA 628
Oy 7192 GAGGCGGACGAGCTTCTCTCAACCGCGCGGAAATCAGGCGATGATATATCCCG 7251
Db 627 GAGGCGGACGAGCTTCTCTCAACCGGTAGACAGCATTTGTGATGCTCA 568
Oy 7252 GGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7311
Db 567 GGGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
Oy 7312 GTATCATTAACCAACAGCGCGGTAGACAGCATTTGTCAAGGCGTCCCTTACAAAG 7371
Db 507 GTATCATTAACCAACAGCGCGGTAGACAGCATTTGTCAAGGCGTCCCTTACAAAG 448

```

```

Oy 7372 GTTTGATCAGCATCTCTCTGCTGATGACCGGCTGTGATGATGATGATGATGAT 7431
Db 447 GCATGCGCTTATCTCATGACCATTTAAATGAATGATGATGATGATGATGATGAT 388
Oy 7432 GATTTGCGTTTGTCTATCTGCTGATGATGATGATGATGATGATGATGATGATGAT 7490
Db 387 GCGGAGAC---CAATTTGTTTAAAGACGCTTGAATGATGATGATGATGATGATGAT 332
Oy 7491 GCAAGAGCTCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7550
Db 331 GCAAGAGCTCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 272
Oy 7551 AGCGCCCGTTTATCCGATGCGCTCTTTGGGACATGCCCCGCTTTGTCCATTTGA 7610
Db 271 GACACCCCTTACGTTCCCAATGGGTGTCTTTGGCCACATTTGCCCTTGTGACATAG 212
Oy 7611 TCTGTCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7670
Db 211 TCACTATGCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 152
Oy 7671 CCCACAAAGATTTTGAAGAAATAGATA 7700
Db 151 CCCACAAAGATTTTGAAGAAATAGATA 122

RESULT 15
CC696106/c 934 bp DNA linear GSS 19-JUN-2003
LOCUS OG10151TV.ZM.0.7.1.5.Zea.mays.genomic.clone.ZMWBMA0448D05,
DEFINITION genomic survey sequence.
ACCESSION CC696106
VERSION CC696106.1 GI:32100882
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 934)
AUTHORS White, J., Quackenbush, J., Van Aken, S., Uterback, T.,
Reinick, A., Fraser, C.M., Budiman, M., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunez, A., Robbins, D. and Lakey, N.
Maize Genomics Consortium
Unpublished (2002)
JOURNAL Other_GSSs: OG10151TV
COMMENT Contact: Cathy White, cwhite@maizegenomics.org
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white@maizegenomics.org
Seq primer: TF
Class: sheared ends.
FEATURES
source location/Qualifiers
1..934
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMWBMA0448D05"
/clone_1lb="ZM.0.7.1.5.KB"
/note="Vector: pCR4-TOPO; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 2.9%; Score 316.6; DB 9; Length 934;
Best Local Similarity 71.4%; Pred. No. 2.4e-53;
Matches 445; Conservative 0; Mismatches 174; Indels 4; Gaps 2;
Oy 1467 AAAAATAAATAAACAAGCGCGGTAGACAGCATTTGTCAAGGCGTCCCTTACAAAG 1526
Db 620 AAAAATAAATAAACAAGCGCGGTAGACAGCATTTGTCAAGGCGTCCCTTACAAAG 561

```

```

Qy 1527 CCGCCGACGACGCTTCTTCTCACCAGCCGCGGAGACCGAGGCGATCGAGTCATATCCCG 1586
Db 560 CGAGGACGACGACCTTCTTCTCACCAGCCGCGGAGACCGAGGCGATCGAGTCATATCCCG 501
Qy 1587 TCGTGGCCGACGACCGCCGCGGAGACCGATCTGTCTCCCGAGGCGATCCCAACTACG 1646
Db 500 TCGTGGCCGAG---CGGCGACCGACCATCTGTCTCCCGAGGCGCGGCTATCCCAACTACG 444
Qy 1647 AGCGCGCGCGCGCTTCAACAGGCTGAGAGTCCGCGATTTCCGACCTCATCCCGCAAG 1706
Db 443 AGCGCGCGCGCGCGCTTCAACAGGCTGAGAGTCCGCGATTTCCGACCTCATCCCGCAAG 384
Qy 1707 GGTGGAGATCGACATCGATCGCTGGAATCCATCCCGCAAGAACACACCGCGCATG 1766
Db 383 GGTGGAGATCGACATCGATCGCTGGAATCCATCCCGCAAGAACACACCGCGCATG 324
Qy 1767 TCATCATTAACCCCAACACCGCGCGGAGGCTTATCTCTACGACCATCTGTCCAG 1826
Db 323 TCATCATTAACCCCAACACCGCGCGGAGGCTTATCTCTACGACCATCTGTCCAG 264
Qy 1827 TTTCAATCTTCTGCTGCTGATTAATGATTCAGTTCAGTGCACCTGTGAATCTTT 1886
Db 263 CACGCTGACCTGCTTATCTTTCGCTGCGCATTTGATGATGAGATCGATCGATTG 204
Qy 1887 TCCCATCGCATATCGATCTGATGTTG-CTCAATTAGTCCGAGGTCGGAAGGCTC 1945
Db 203 TGTGTTTAAAGACACTGCTGCTGTCTGTCTATCTAGTTCGCGAGTGCAGAGACTT 144
Qy 1946 GGAATATTTGATTTGCTGACGAGATTAACGCGAACGCTGTTCTGCGGACGCGCGCTTC 2005
Db 143 GGAATATTTGATTTGCTGACGAGATTAACGCGAACGCTGTTCTGCGGACGCGCGCTTC 84
Qy 2006 ATCCCAATGGAGTGTGGGACATCAACCGCTGCTGCTGATGAGGCTCTGTCACG 2065
Db 83 GTCCCATGAGGCTGCTTATCTTTCGCTGCGCATTTGATGATGAGATCGATCGATTG 24
Qy 2066 TCATGATATGTCCTGATGATGCG 2088
Db 23 AGATGATATGTCCTGATGATGCG 1

```

Search completed: October 15, 2005, 11:53:19
 Job time : 30434.4 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 13:04:36 ; Search time 66.582 Seconds
 (without alignment) 13720.565 Million cell updates/sec

Title: US-10-019-783-3
 Perfect score: 24
 Sequence: 1 tcgsgatcagtcgacacaca 24

Scoring table:
 IDENTITY_NUC
 Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
 Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_g881:*

9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.4	85.0	999	4	BG334930
2	19.8	82.5	563	1	AJ807947
3	19.4	80.8	723	5	BU612123
4	19.4	80.8	946	5	BU505136
5	19.2	80.0	198	1	AA359060
6	19.2	80.0	224	1	AA359061
7	19.2	80.0	279	7	CN495808
8	19.2	80.0	478	7	CN495808
9	18.8	78.3	391	2	BE715690
10	18.8	78.3	544	5	BQ015926
11	18.8	78.3	541	2	BE706168
12	18.8	78.3	579	8	B05889
13	18.8	78.3	654	4	BG334045
14	18.8	78.3	926	5	BX371803
15	18.8	78.3	940	5	BX371804
16	18.8	78.3	989	5	BQ068921
17	18.8	78.3	1004	5	BQ069095
18	18.8	78.3	1190	8	CC242470
19	18.8	78.3	1403	3	HSMB0511
20	18.8	78.3	8864	3	HSMB08880
21	18.4	76.7	532	4	BG996884
22	18.4	76.7	552	8	BZ266969
23	18.4	76.7	571	8	AQ465392
24	18.2	75.8	260	2	BE932400
25	18.2	75.8	442	2	BE509060
26	18.2	75.8	451	8	B64254
27	18.2	75.8	463	7	CY359206
28	18.2	75.8	530	4	BM728080
29	18.2	75.8	596	1	AL791979
30	18.2	75.8	600	7	CO358000
31	18.2	75.8	630	1	AL862290
32	18.2	75.8	649	1	AL859034
33	18.2	75.8	654	1	AL887699
34	18.2	75.8	658	1	AL864891
35	18.2	75.8	661	1	AL875887
36	18.2	75.8	682	4	AL861073
37	18.2	75.8	685	4	BI186061
38	18.2	75.8	688	7	CR433911
39	18.2	75.8	726	5	BX776131
40	18.2	75.8	727	5	BX731133
41	18.2	75.8	750	4	BI183205
42	18.2	75.8	751	5	BX750164
43	18.2	75.8	762	7	CR433912
44	18.2	75.8	792	7	CK127026
45	18.2	75.8	826	5	BX781066

ALIGNMENTS

RESULT 1

LOCUS BG334930/1

DEFINITION 602461255F1 NIH_MGC_20 Homo sapiens CDNA clone IMAGE:4578228 5', 999 bp mRNA linear EST 27-FEB-2001

ACCESSION mRNA sequence.
 VERSION BG334930
 KEYWORDS EST. 13141368
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE 1 (bases 1 to 999)
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DFP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Inceye Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN, at: http://image.llnl.gov
 Plate: L1CM1293 row: m column: 13
 High quality sequence stop: 520.
 Location/Qualifiers
 1..999
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:457828"
 /tissue_type="melanocytic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 20"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
 Query Match 85.0%; Score 20.4; DB 4; Length 999;
 Best Local Similarity 95.5%; Pred. No. 1.1e+02;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 TCTGGATCTCAGTCATCA 22
 |||||
 Db 567 TCTGGATCTCAGTCATCA 546

RESULT 2
 AJ807947 563 bp mRNA linear EST 11-MGC-2004
 LOCUS AJ807947
 DEFINITION Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018_6_08 d24, mRNA sequence.
 ACCESSION AJ807947
 VERSION AJ807947.1 GI:51123275
 KEYWORDS EST.
 SOURCE Antirrhinum majus (snapdragon)
 ORGANISM Antirrhinum majus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Plantaginaceae; Antirrhinae; Antirrhinum.

REFERENCE
 1 (bases 1 to 563)
 AUTHORS Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.
 TITLE Antirrhinum EST Collection
 JOURNAL Unpublished (2003)
 COMMENT Contact: Schwarz-Sommer Z
 Molekulare Pflanzen-genetik
 MPI fuer Zuechtungs-forschung
 Carl-von-Linne Weg 10, D-50829, Germany.

FEATURES
 source Location/Qualifiers
 1..563
 /organism="Antirrhinum majus"
 /mol_type="mRNA"
 /db_xref="taxon:4151"
 /clone="018_6_08 d24"
 /tissue_type="whole plant"
 /clone_1lb="Antirrhinum majus whole plant"

ORIGIN
 Query Match 82.5%; Score 19.8; DB 1; Length 563;
 Best Local Similarity 91.3%; Pred. No. 1.9e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 CTGGATCTCAGTCATCA 24
 |||||
 Db 233 CTGGATCTCAGTCATCA 255

RESULT 3
 BU612123 723 bp mRNA linear EST 20-FEB-2003
 LOCUS BU612123
 DEFINITION UT-M-EMO-cax-k-01-0-UI r1 NIH BMAP EMO Mus musculus CDNA clone UT-M-EMO-cax-k-01-0-UI 5', mRNA sequence.
 ACCESSION BU612123
 VERSION BU612123.1 GI:23278338
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus.
 1 (bases 1 to 723)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..723
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="UT-M-EMO-cax-k-01-0-UI"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="NIH BMAP EMO"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 80.8% Score 19.4; DB 5; Length 723;
 Best Local Similarity 95.2%; Pred. No. 36+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTCATCCACA 24
 |||||
 DB 475 GGGATCTCAGTCATCCACA 495

RESULT 4
 B0505136 946 bp mRNA linear EST 12-SEP-2002
 LOCUS AGENCOURT 10013187 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6493080
 DEFINITION 5', mRNA sequence.
 ACCESSION B0505136
 VERSION B0505136.1 GI:22811369
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE NIH-MGC http://mgi.mgi.nhl.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strusberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: The Cepho Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M14047 row: c column: 01
 High quality sequence stop: 628.
 Location/Qualifiers

FEATURES
 source 1..946
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6493080"
 /ciseq_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /clone_11b="NIH_MGC_94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 80.8% Score 19.4; DB 5; Length 946;
 Best Local Similarity 95.2%; Pred. No. 3.1e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTCATCCACA 24
 |||||
 DB 689 GGGATCTCAGTCATCCACA 709

RESULT 5
 AA359060/c 198 bp mRNA linear EST 21-APR-1997
 LOCUS EST68008 Fetal lung II Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA359060
 VERSION AA359060.1 GI:2011377
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 198)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bul,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man'hai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fitchman,J.L., Geoghagen,N.S.,
 Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Seadek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunach,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
 Raymond,L., Wei,X.F., Wang,J., Xu,C., Yu,G.H., Ruben,S.M.,
 Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Frazer,C.M., and Venter,J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 96026280
 7566098

TITLE

JOURNAL

PUBMED

COMMENT

Other ESTs: TH076983
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (http://www.tigr.org/cdb/hgi/hgi.html)
 Seq primer: M13 Reverse.
 Location/Qualifiers

FEATURES

source 1..198
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):163152"
 /db_xref="taxon:9606"
 /sex="male"
 /dev stage="fetus, 19 wks"
 /clone_11b="Fetal lung II"
 /note="Organ: lung; Vector: pBlueScript SK-; Site_1:
 EcoRI; Site_2: XhoI"

ORIGIN

Query Match 80.0% Score 19.2; DB 1; Length 198;
 Best Local Similarity 87.5%; Pred. No. 3.2e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCTGGATCTCAGTCATCCACA 24
 |||||
 DB 162 TCTGGATCTCAGTCATCCACA 139

RESULT 6
 AA359061/c 224 bp mRNA linear EST 21-APR-1997
 LOCUS EST68009 Fetal lung II Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA359061
 VERSION AA359061.1 GI:2011378
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 224)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bul,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man'hai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,

Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodex, A., Gnehm, C.L., Hanna, M.C., Hedblom, B., Hinkle, P.S.Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.B., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Korak, D.L., Kunach, C., Hungjun, J., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

JOURNAL
MEDLINE
96026280
7566098
Other ESTs: TH076983
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/cdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source
1..224
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):163153"
/db_xref="taxon:9606"
/sex="male"
/dev_stage="fetus, 19 wks"
/clone_lib="Fetal lung II"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 80.0%; Score 19.2; DB 1; Length 224;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OR 1 TCTGGGATCTCAGTCATCCACCA 24
|||||
Db 162 TCTGGGCTTTCAGTCATCCATCA 139

RESULT 7
CN495808/c 279 bp mRNA linear EST 24-MAY-2004
LOCUS Mdfw2020p20.y1 Mdfw Malus x domestica cDNA clone Mdfw2020p20 5' DEFINITION similar to TR:Q9ZP29 Q9ZP29 PUTATIVE GRI-1-LIKE PROTEIN. ; mRNA sequence.
CN495808
ACCESSION CN495808.1 GI:46597534
VERSION
KEYWORDS EST.
SOURCE Malus x domestica (cultivated apple)
ORGANISM Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosida I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 279)
Korban, S., Vockin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A., Alwinckle, H., Malinov, M., Carroll, N., Goldsbrough, P., Orysi, K., Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritzer, E., Ronko, I., Tesgarestvill, R., Kennedy, S., Waterston, R., and Wilson, R.
Apple Functional Genomics grant - NSF 0321702

TITLE

JOURNAL
COMMENT
Unpublished (2004)
Contact: Schuyler S. Korban
Apple Functional Genomics grant - NSF 0321702
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: esat@watson.wustl.edu
Library materials provided by: Schuyler S. Korban Library
constructed by: A. Hernandez / K. Gasic Library sequenced by:
Washington University Genome Sequencing Center
WashU EST name: aa180h10.y1
High quality sequence stop: 255.

FEATURES
source
1..279
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="Mdfw2020p20"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Mdfw"
/note="Vector: DH10B ampicillin resistant; Site_1: NotI; Site_2: EcoRI; Total RNA was extracted separately from each stage (bud, balloon, open and after pollination) using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex direct RNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo18(dT) primer with an identifying tag sequence (see table below). cDNAs from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (bud) insert 18(A)TGCAG; Stage 2 (balloon) insert 18(A)TGCAG; Stage 3 (open) insert 18(A)TGCAG; Stage 4 (after pollination) insert 18(A)TGCAG; Tag identification when sequencing from 3' end: Stage 1 (bud) TCCGAl8(T) insert; Stage 2 (balloon) TCCGAl8(T) insert; Stage 3 (open) ACCGAl8(T) insert; Stage 4 (after pollination) ACCGAl8(T) insert. Double stranded cDNAs were size selected (more than 450 bp), adapted with EcoRI adapters at both ends and then digested with NotI. The cDNAs were then directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene). Identification of adaptors and tags in 5'-end sequenced clones: <Vector>...TAGCTT<End Vector><Start EcoRI adaptor>GATTCGAATTCATGCTGTGGG<End EcoRI adaptor><Start insert>...AAAAAAAAAAAAAAAA<End insert><Start Tag>TGCAG<End Tag><Start NotI site>Vector>GCCGCGCCACCGCGG... The total number of white colony forming units (cfu) in the primary library before amplification was 1.1x10⁶ cfu (colony forming units). The background of empty clones was less than 1%. Inserts ranged from 0.5kb to 3 kb, as determined by PCR. Purified plasmid DNA from the primary library was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 9x10⁶ cfu. Background of empty clones was less than 1%.

ORIGIN

Query Match 80.0%; Score 19.2; DB 7; Length 279;
Best Local Similarity 87.5%; Pred. No. 3.3e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OR 1 TCTGGGATCTCAGTCATCCACCA 24
|||||

Db 197 TCTGTGATCTGAGTCATCCACAGA 174

RESULT 8
LOCUS A0880175 478 bp DNA 1linear GSS 09-NOV-1999
DEFINITION HS 5037 B1.D11.T7 RPCR-11 Human Male BAC Library Homo sapiens genomic clone Plate=8805 Col=21 Row=H, genomic survey sequence.
ACCESSION A0880175
KEYWORDS A0880175.1 GI:6311642
SOURCE GSS.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 478)
AUTHORS Mahatras,G.G., Wallace,J.C., Satch,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
10449764
JOURNAL
MEDLINE
PubMed
COMMENT Contact: Mahatras GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCR-11. For BAC library availability, please contact Plietier de Jong (plietier@editions.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (<http://www.resgen.com>). BAC end web server: <http://www.htec.washington.edu>
Plate: 8805 row: H column: 21
Seq primer: T7
Class: BAC ends
High quality sequence stop: 478.
Location/Qualifiers
1..478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=8805 Col=21 Row=H"
/sex="male"
/clone_1ib="RPCR-11 Human Male BAC Library"
/note="Vector: PBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the PBAC3.6 vector at EcoRI sites"

ORIGIN

Query Match 80.0%; Score 19.2; DB 8; Length 478;
Best Local Similarity 87.5%; Pred. No. 3.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 TCTGGATCTCAGTCATCCACACA 24
|||||
313 TCTGGAGCTCATTCATCTCTACA 336
|||||

Db 313 TCTGGAGCTCATTCATCTCTACA 336
|||||

RESULT 9
LOCUS BE715690/c 391 bp mRNA 1linear EST 12-SEP-2000
DEFINITION CM2-HT0750-040700-250-cl1 HT0750 Homo sapiens CDNA, mRNA sequence.
ACCESSION BE715690
KEYWORDS BE715690.1 GI:10103955
EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongseneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
JOURNAL
MEDLINE
PubMed
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?rl=et2=CM2-HT0750-040700-250-cl1&t3=2000-07-04&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 391.
Location/Qualifiers
1..391
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1ib="HT0750"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORFESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 78.3%; Score 18.8; DB 2; Length 391;
Best Local Similarity 90.9%; Pred. No. 5.4e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 CTGGATCTCAGTCATCCAC 23
|||||
228 CTGGGTTCTCAGTCATCCAGC 207
|||||

Db 228 CTGGGTTCTCAGTCATCCAGC 207
|||||

RESULT 10
LOCUS B0015926 444 bp mRNA 1linear EST 17-JUN-2002
DEFINITION UI-H-DT1-8v2-d-05-0-UI-81 NCI_CGAP_DTI Homo sapiens CDNA clone IMAGE:5886292 3', mRNA sequence.
ACCESSION B0015926
KEYWORDS B0015926.1 GI:19751203
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/INM at: <http://image.llnl.gov>
 Seq primer: M13 FORWARD
 POLYA=Yes

FEATURES

source

Location/Qualifiers
 1..444
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5886292"
 /tissue_type="Metastatic Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_id="NCI_CGAP_D11"
 /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 NCI CGAP D11 is a normalized cDNA library containing the
 following tissue(s): Metastatic Chondrosarcoma in Lung.
 The library was constructed according to Bonaldo, Lennon
 and Soares, Genome Research, 6:791-806, 1996. First strand
 cDNA synthesis was primed with an oligo-dT primer
 containing a Not I site. Double stranded cDNA was ligated
 to an EcoR I adaptor, digested with Not I, and cloned
 directionally into pT7T3-Pac vector. The oligonucleotide
 used to prime the synthesis of first-strand cDNA contains
 a library tag sequence that is located between the Not I
 site and the (dT)18 tail. The sequence tag for this
 library is AACTGTCGC.
 TAG_TISSUE=Lung metastatic chondrosarcoma
 TAG_LIB=UI-H-D11
 TAG_SEQ=AACTGTCGC"

ORIGIN

Query Match 78.3%; Score 18.8; DB 5; Length 444;
 Best Local Similarity 90.9%; Pred. No. 5.5e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 2 CTGGGATCTCAGTCATCAAC 23
 400 CTGGGATCTCAGTCATCAAC 421

RESULT 11
 BE706168/c 541 bp mRNA linear EST 12-SEP-2000
 LOCUS RCI-HT0229-160600-112-f07 HT0229 Homo sapiens CDNA, mRNA sequence.
 DEFINITION BE706168
 ACCESSION BE706168.1 GI:10094433
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 541)
 Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,
 Nagai M.A., da Silva W.J., Zago M.A., Bordin S., Costa F.F.,
 Goldman G.H., Carvalho A.F., Matukuma A., Bala G.S., Simpson D.H.,
 Bruneisen A., de Oliveira P.S., Bucher P., Jongeneel C.V.,
 O'Hare M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and
 Simpson A.J.

REFERENCE

Shotgun sequencing of the human transcriptome with ORF expressed

TITLE

Sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research

FEATURES

source

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: atempone@ludwig.org.br
 This sequence was derived from the PABSP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RCI-HT0229-160>
 600-112-f07et3-2000-06-16et4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 541.

Location/Qualifiers
 1..541
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_id="HT0229"
 /note="Organ: head, neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 78.3%; Score 18.8; DB 2; Length 541;
 Best Local Similarity 90.9%; Pred. No. 5.6e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 2 CTGGGATCTCAGTCATCAAC 23
 417 CTGGGATCTCAGTCATCAAC 396

RESULT 12
 B05889 579 bp DNA linear GSS 13-JUL-1996
 LOCUS CSR1-72911-u CSR1 flow sorted chromosome 11 specific cosmid Homo
 DEFINITION B05889
 ACCESSION B05889.1 GI:1415167
 VERSION GSS.
 KEYWORDS Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 579)
 Evans G.A., Burbee D., Davies C., Hahner L., Oliver T., Gilbert M.,
 Jones D., Ward T., Gillilan E., Schagemann J., Probst S.,
 Harris J., Deford J., McFarland J., Burzinski K., Khan M.,
 Kupfer K. and Garner H.R.
 Genomic Sequence Sampled Map of Chromosome 11
 Unpublished (1996)
 Contact: Evans GA, Shane Probst
 McDermott Center for Human Growth and Development
 University of Texas Southwestern Medical Center At Dallas
 5323 Harry Hines Blvd, Dallas TX 75235-8591
 Tel: 214-648-1600
 Fax: 214-648-1666
 Email: gevans@utsw.bmed.edu, shane@mcdermott.bmed.edu
 PCR Primers
 FORWARD: TGTTGAGAAAGTTCAGAAC
 BACKWARD: CCATGCTAGAACATCATC
 Seq primer: T7
 Class: cosmid ends
 High quality sequence stop: 579.

TITLE

Location/Qualifiers
 1..579
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES

source

/clone="CSRL-72g11"
 /sex="female"
 /cell_type="chimeric hamster somatic cell hybrid"
 /clone_lib="CSRL flow sorted Chromosome 11 specific cosmid"
 /note="Vector: sCos-1; Human Chromosome 11 specific cosmid library prepared from flow sorted human Chromosome 11 derived from Chinese Hamster Ovary (CHO) monochromosomal somatic cell hybrid, J1"

ORIGIN

Query Match 78.3%; Score 18.8; DB 8; Length 579;
 Best Local Similarity 90.9%; Pred. No. 5.7e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGGATCTCAGTCATCCAC 23
 DB 239 CTGGGTTCTCAGTCATCCAGC 218

RESULT 13
 BG334045/C 654 bp mRNA linear EST 27-FEB-2001

LOCUS 602460411P1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4577258 5',
 DEFINITION mRNA sequence.

ACCESSION

BG334045

VERSION

BG334045.1 GI:13140483

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 654)
 NIH-MGC http://mgi.nci.nih.gov/
 Unpublished (1993)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DC/DTP

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CDNA library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LCM1291 row: e column: 03
 High quality sequence stop: 652.
 Location/Qualifiers

FEATURES

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4577258"
 /cfeature="IMAGE:4577258"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 20"
 /note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 78.3%; Score 18.8; DB 4; Length 654;
 Best Local Similarity 90.9%; Pred. No. 5.8e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGGATCTCAGTCATCCAC 23
 DB 560 CTGGGTTCTCAGTCATCCAGC 539

RESULT 14

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-G1190(dt) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 5982.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0BA1024Zf07_CS02262_1&c=5982.f

FEATURES

source

Location/Qualifiers
 1..926
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1061Y108"
 /cfeature="CS0D1061Y108"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 78.3%; Score 18.8; DB 5; Length 926;
 Best Local Similarity 90.9%; Pred. No. 6.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGGATCTCAGTCATCCAC 23
 DB 503 CTGGGTTCTCAGTCATCCAGC 524

RESULT 15

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-G1190(dt) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 5982.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0BA1024Zf07_CS02262_1&c=5982.f

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 5982.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BA1024ZF07_CS02262_2&c=5982.f

FEATURES

Source

Location/Qualifiers
 1. 940
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1061Y108"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 78.3%; Score 18.8; DB 5; Length 940;
 Best Local Similarity 90.9%; Pred. No. 6.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGGGATCTCAGTCATCCAC 23
 |||||
 Db 507 CTGGGTCTCAGTCATCCAGC 528

Search completed: October 15, 2005, 11:53:26
 Job time : 73.582 secs

CC vigorous growth in calcareous alkaline soil for improving crop
CC production. The constructed plant has tolerance to iron deficiency, and
CC is therefore capable of vigorous growth in calcareous alkaline soil. The
CC present sequence encodes two nicotianamine aminotransferases (NAT),
CC designated NAT-A and NAT-B, isolated from *Hordeum vulgare* L. var. Igri
CC (barley), for use in the method of the invention

XX Sequence 10966 Bf; 3082 A; 2422 C; 2366 G; 3096 T; 0 U; 0 Other;

Query Match 100.0%; Score 10966; DB 5; Length 10966;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 10966; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGATCCCATTTGCAATGGATGATTTAGCTATCAACGAAAGAAAGAGATGGCATATGCC 60
DB 1 CTCGATCCCATTTGCAATGGATGATTTAGCTATCAACGAAAGAAAGAGATGGCATATGCC 60

QY 61 CTGTGTGTATCTCTCACTGCTGCTGCGAATGGCGATACCGAGTTAGGTAGTGTCTTTT 120
DB 61 CTGTGTGTATCTCTCACTGCTGCTGCGAATGGCGATACCGAGTTAGGTAGTGTCTTTT 120

QY 121 TTAGCATGATGTCTGCGGCACTGCGCAAGAAACTGCGTGCAGCGGAGCTGCAGAGATT 180
DB 121 TTAGCATGATGTCTGCGGCACTGCGCAAGAAACTGCGTGCAGCGGAGCTGCAGAGATT 180

QY 181 GAGCGATGATGCTTTGTGATGAGCGGAGCTGAGTGGGTGTCTACTACTGAAACCCATCA 240
DB 181 GAGCGATGATGCTTTGTGATGAGCGGAGCTGAGTGGGTGTCTACTACTGAAACCCATCA 240

QY 241 GCATTGGGTGATGCGATGCGAAGCATATCTCTTCTGCGTCCGATCCGTTATCTTT 300
DB 241 GCATTGGGTGATGCGATGCGAAGCATATCTCTTCTGCGTCCGATCCGTTATCTTT 300

QY 301 TTTCTCCAAATTAATTAAGAGGATAGATGATGATGATGATGATGATGATGATGATGATG 360
DB 301 TTTCTCCAAATTAATTAAGAGGATAGATGATGATGATGATGATGATGATGATGATGATG 360

QY 361 ATAGAACCAAGAGCGCGCGCAAAATATGCGGAGATATATGATGAGCGCGCAACCGC 420
DB 361 ATAGAACCAAGAGCGCGCGCAAAATATGCGGAGATATATGATGAGCGCGCAACCGC 420

QY 421 CACGCGCGTCACTGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 480
DB 421 CACGCGCGTCACTGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 480

QY 481 TCTTCTGCT 540
DB 481 TCTTCTGCT 540

QY 541 ACCGCTACTCAATCCGTCGCCCACTGTTTCTTATTAACCGCTACTAGTATTCCTGCT 600
DB 541 ACCGCTACTCAATCCGTCGCCCACTGTTTCTTATTAACCGCTACTAGTATTCCTGCT 600

QY 601 GATAGCTAGTATCT 660
DB 601 GATAGCTAGTATCT 660

QY 661 ACGCGAGAGCGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACGCGAGAGCGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720

QY 721 GAGCAACG 780
DB 721 GAGCAACG 780

QY 781 TGCGCAACG 840
DB 781 TGCGCAACG 840

QY 841 CAACG 900
DB 841 CAACG 900

QY 901 CAACG 960
DB 901 CAACG 960

QY 961 CGAGCATGCGCGAGGATCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 961 CGAGCATGCGCGAGGATCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020

QY 1021 AGAGAGAGAGAGCGCGGTGAGTGAATTTTCCGCGGTGCCAAGGACCGCGTGTGCGCGCG 1080
DB 1021 AGAGAGAGAGAGCGCGGTGAGTGAATTTTCCGCGGTGCCAAGGACCGCGTGTGCGCGCG 1080

QY 1081 GACCGGCGCGCAATGAGCATCCGCGCGATCGGTATCAAGATCAAGCGCGAGCGTGCAGGA 1140
DB 1081 GACCGGCGCGCAATGAGCATCCGCGCGATCGGTATCAAGATCAAGCGCGAGCGTGCAGGA 1140

QY 1141 GAAAGGCTT 1200
DB 1141 GAAAGGCTT 1200

QY 1201 CCGCAACGCGCGTGCAGGCGCGAGAGCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCTT 1260
DB 1201 CCGCAACGCGCGTGCAGGCGCGAGAGCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCTT 1260

QY 1261 CTGCTTACCCCGCGCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1320
DB 1261 CTGCTTACCCCGCGCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1320

QY 1321 GAACAATTTCTTTTCTGCGGTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1380
DB 1321 GAACAATTTCTTTTCTGCGGTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1380

QY 1381 TTTCACTGTCCGCGCGTGCAGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
DB 1381 TTTCACTGTCCGCGCGTGCAGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440

QY 1441 GACTTCTTCCCACTTATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500
DB 1441 GACTTCTTCCCACTTATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1500

QY 1501 GTGCGAGGCGTGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB 1501 GTGCGAGGCGTGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560

QY 1561 CCAAGCGATGAGGATCAATATCCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1620
DB 1561 CCAAGCGATGAGGATCAATATCCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1620

QY 1621 CCGCAAGCGCGTATCCCAACCTACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
DB 1621 CCGCAAGCGCGTATCCCAACCTACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680

QY 1681 GCATTTGCACTCATCCCGCAAGAGGCGTGGAGATGACATGACATGACATGACATGACATG 1740
DB 1681 GCATTTGCACTCATCCCGCAAGAGGCGTGGAGATGACATGACATGACATGACATGACATG 1740

QY 1741 CGCGCAAGAGCAACACCGCGCATGTCATATAACCCCAACACCGCGCGCGCGCGCGCGCT 1800
DB 1741 CGCGCAAGAGCAACACCGCGCATGTCATATAACCCCAACACCGCGCGCGCGCGCGCGCT 1800

QY 1801 TTAATCTCTAGCAACATCTGTCCAAAGTTTCAATCTTTTGGCTTGTGCAATATGATTTCA 1860
DB 1801 TTAATCTCTAGCAACATCTGTCCAAAGTTTCAATCTTTTGGCTTGTGCAATATGATTTCA 1860

QY 1861 GTTCAAGTCACTCTGTGATTTCTTTTGGCAATGCAATGCAATGCAATGCAATGCAATGCA 1920
DB 1861 GTTCAAGTCACTCTGTGATTTCTTTTGGCAATGCAATGCAATGCAATGCAATGCAATGCA 1920

QY 1921 GGTGCGGAGGTGCGAAAGGCTCGAATATGATGATGATGATGATGATGATGATGATGATG 1980
DB 1921 GGTGCGGAGGTGCGAAAGGCTCGAATATGATGATGATGATGATGATGATGATGATGATG 1980

QY 1981 GCTGCTTCTGGGCAAGCGCGCGCTTCTATCCCAATGGAGTGTGTTGGGCAATCAACCCCTG 2040

Db 1981 GCTGGTTCGGGACGCCCCGTTTCATCCCATGGGAGGTTTGGGACATCACCCCTGT 2040
Qy 2041 GCTGTCATAGGGCTCTGTCCAGTCATGATAGTGCCTGATGGCGGCTTGATGGGT 2100
Db 2041 GCTGTCATAGGGCTCTGTCCAGTCATGATAGTGCCTGATGGCGGCTTGATGGGT 2100
Qy 2101 AGCGGATGACGACCCGAGAAATCTTACAGAACTAGGATCTTAAATCTCTATATCA 2160
Db 2101 AGCGGATGACGACCCGAGAAATCTTACAGAACTAGGATCTTAAATCTCTATATCA 2160
Qy 2161 TTCTTTTCAATGCTACTAGGATGATTAATTAAGTACTAGTCAATATATTTGGCTAAT 2220
Db 2161 TTCTTTTCAATGCTACTAGGATGATTAATTAAGTACTAGTCAATATATTTGGCTAAT 2220
Qy 2221 TTGATCTGACATTTTGTGTGATCTCTACATCAATTAAGAAATCTCAATGTCTGCA 2280
Db 2221 TTGATCTGACATTTTGTGTGATCTCTACATCAATTAAGAAATCTCAATGTCTGCA 2280
Qy 2281 CAGACCCGACACCTTCACTAGGTCAGTCTTGGATTAATCTGTTTCAAGAAATTA 2340
Db 2281 CAGACCCGACACCTTCACTAGGTCAGTCTTGGATTAATCTGTTTCAAGAAATTA 2340
Qy 2341 GTCTTTGGTATTTACTCTCTCTGTCCTATTTTGTGCTCCGCTCCCTATGTTGAGGACGC 2400
Db 2341 GTCTTTGGTATTTACTCTCTCTGTCCTATTTTGTGCTCCGCTCCCTATGTTGAGGACGC 2400
Qy 2401 CAGTCGATGTCAGTACGCGTTTTTTCATTAAGTTGAAGTCAAGTCAACAT 2460
Db 2401 CAGTCGATGTCAGTACGCGTTTTTTCATTAAGTTGAAGTCAAGTCAACAT 2460
Qy 2461 ACACCTTGATGATTTTACTCTTGTGTTGTTGATCCGATTAATAAATAAATAAATAA 2520
Db 2461 ACACCTTGATGATTTTACTCTTGTGTTGTTGATCCGATTAATAAATAAATAAATAA 2520
Qy 2521 CTGAACCTACTGTGATTAACAACCTGTTCTTAACAAGATACATGATGACTATGGG 2580
Db 2521 CTGAACCTACTGTGATTAACAACCTGTTCTTAACAAGATACATGATGACTATGGG 2580
Qy 2581 CATGCCATATCTTTTGGGTCAAGTATGAGTATGTTGGACCTCTTTTGAATAATAGT 2640
Db 2581 CATGCCATATCTTTTGGGTCAAGTATGAGTATGTTGGACCTCTTTTGAATAATAGT 2640
Qy 2641 ACATGTAATGATGATTAACAATTTTATTAAGAATTTCAATTTGATATCTTGTATGTT 2700
Db 2641 ACATGTAATGATGATTAACAATTTTATTAAGAATTTCAATTTGATATCTTGTATGTT 2700
Qy 2701 ATTGTTCTGTGATTCACAGATTTACTTGGGTTTTTGTGATCAATCAATTTGTTGAG 2760
Db 2701 ATTGTTCTGTGATTCACAGATTTACTTGGGTTTTTGTGATCAATCAATTTGTTGAG 2760
Qy 2761 GCAAGCTCTTCTCAGATTTCTTGAGAACACAAAGAAATTTCTTTAAGGCGATTAATGGT 2820
Db 2761 GCAAGCTCTTCTCAGATTTCTTGAGAACACAAAGAAATTTCTTTAAGGCGATTAATGGT 2820
Qy 2821 CTGCTTAAAGGATTCATCAGAGTATGCTTCAAAACAATTAAGAAACAATTAACATTA 2880
Db 2821 CTGCTTAAAGGATTCATCAGAGTATGCTTCAAAACAATTAAGAAACAATTAACATTA 2880
Qy 2881 TGTCTTCAAGGACGAGAGATCAATGTTGTCTAGTGAAGCTATTTTGTGAAGTAA 2940
Db 2881 TGTCTTCAAGGACGAGAGATCAATGTTGTCTAGTGAAGCTATTTTGTGAAGTAA 2940
Qy 2941 AAATCTTGAAGAGTCTCAGTAAATCAATTAATTAATTAAGATTAATCTGGACCGAA 3000
Db 2941 AAATCTTGAAGAGTCTCAGTAAATCAATTAATTAATTAAGATTAATCTGGACCGAA 3000
Qy 3001 ATGCAATCCAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Db 3001 ATGCAATCCAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Qy 3061 TGAAGATGATGATTAATTAAGATTAAGAGAGCTTAAAGTTATGCTACGATATCAT 3120
Db 3061 TGAAGATGATGATTAATTAAGATTAAGAGAGCTTAAAGTTATGCTACGATATCAT 3120

Db 3061 TGAAGATGATGATTAATTAAGATTAAGAGAGCTTAAAGTTATGCTACGATATCAT 3120
Qy 3121 CTGATATATCTTTTGTCCATCTTTTGTATTAATCTGAGAAATGTTATGTTCAATCATATT 3180
Db 3121 CTGATATATCTTTTGTCCATCTTTTGTATTAATCTGAGAAATGTTATGTTCAATCATATT 3180
Qy 3181 ACACAAATATCAATGCTATGTTTCTAGAAAGATGATTAATTTTCTGTAACATGAACCTC 3240
Db 3181 ACACAAATATCAATGCTATGTTTCTAGAAAGATGATTAATTTTCTGTAACATGAACCTC 3240
Qy 3241 GTATTAATCTCAGTAAACAGGTAAACCTGAATCAATCTTTTGAAGAAATAGACA 3300
Db 3241 GTATTAATCTCAGTAAACAGGTAAACCTGAATCAATCTTTTGAAGAAATAGACA 3300
Qy 3301 TGAATTAATCTTCAATGTAACAGGTAAACCTGAATCAATCTTTTGAAGAAATAGACA 3360
Db 3301 TGAATTAATCTTCAATGTAACAGGTAAACCTGAATCAATCTTTTGAAGAAATAGACA 3360
Qy 3361 GAAATCATTGTTGATTTTGTGACTGTAATGAAGTCTTATCAATTTCCGAGATGACTATA 3420
Db 3361 GAAATCATTGTTGATTTTGTGACTGTAATGAAGTCTTATCAATTTCCGAGATGACTATA 3420
Qy 3421 CATATTAATGATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Db 3421 CATATTAATGATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Qy 3481 AAGACTTTTAAACAATATTTCTGTAAGTCAATATCAATGATCTCTCTGTTTTC 3540
Db 3481 AAGACTTTTAAACAATATTTCTGTAAGTCAATATCAATGATCTCTCTGTTTTC 3540
Qy 3541 TTGACCTGATGCTCTGTCACATGTTGATGCTCCACAAATAATAGAACACATGTT 3600
Db 3541 TTGACCTGATGCTCTGTCACATGTTGATGCTCCACAAATAATAGAACACATGTT 3600
Qy 3601 TCAATTAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660
Db 3601 TCAATTAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660
Qy 3661 TTTTGTAGTCCAACTTAATTAATTTGTGACTATTTCCAAAACAATTAATTAATTAAT 3720
Db 3661 TTTTGTAGTCCAACTTAATTAATTTGTGACTATTTCCAAAACAATTAATTAATTAAT 3720
Qy 3721 CCTCTCAATGATGAGGAGTCTTGAATGAGCAATGAGGCTCCGATTAATCTTTGCTTG 3780
Db 3721 CCTCTCAATGATGAGGAGTCTTGAATGAGCAATGAGGCTCCGATTAATCTTTGCTTG 3780
Qy 3781 TGTTCATCTTCTCTTCAAGATGCTCGGAAGATCAATCATCTGTCAAAGAACAA 3840
Db 3781 TGTTCATCTTCTCTTCAAGATGCTCGGAAGATCAATCATCTGTCAAAGAACAA 3840
Qy 3841 GAAAGAAATTCGACGATGATGCTAGTGTATATCTGACTGAAGCTGTAATCATTC 3900
Db 3841 GAAAGAAATTCGACGATGATGCTAGTGTATATCTGACTGAAGCTGTAATCATTC 3900
Qy 3901 CAGTATCCCAATCTATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
Db 3901 CAGTATCCCAATCTATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
Qy 3961 CAACATCTCTTGAATATGTTCTGTTGTGTTGCTGAGACGAAACATAGTAATGTTAT 4020
Db 3961 CAACATCTCTTGAATATGTTCTGTTGTGTTGCTGAGACGAAACATAGTAATGTTAT 4020
Qy 4021 GTTATGAAAGTTACATTTGCGCTGAAAGTCTTTGAAGTTTTTTTTTTTTTTGGGGGGG 4080
Db 4021 GTTATGAAAGTTACATTTGCGCTGAAAGTCTTTGAAGTTTTTTTTTTTTTTGGGGGGG 4080
Qy 4081 GGGGGGGGGGCTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
Db 4081 GGGGGGGGGGCTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
Qy 4141 CTTTGTAAACAATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGAT 4200
Db 4141 CTTTGTAAACAATTTGATGATTTTGTATGATGATGATGATGATGATGATGATGATGAT 4200

4201 AAGCAAAACCTTACTTATGTTAGGCTCACTAAGGAGGCTTTGGTTGAGAGAGAGAA 4260
4201 AAGCAAAACCTTACTTATGTTAGGCTCACTAAGGAGGCTTTGGTTGAGAGAGAGAA 4260
4261 GGATCAGTTGATGATATCCCAATCATCGAAGTAAATCATGTGTGTTGCTACCACTTTT 4320
4261 GGATCAGTTGATGATATCCCAATCATCGAAGTAAATCATGTGTGTTGCTACCACTTTT 4320
4321 CTACAAATCTTACTGATGATGAGCTGAGTGAATCTGATCAACACCTGACACCAATATT 4380
4321 CTACAAATCTTACTGATGATGAGCTGAGTGAATCTGATCAACACCTGACACCAATATT 4380
4321 CTACAAATCTTACTGATGATGAGCTGAGTGAATCTGATCAACACCTGACACCAATATT 4380
4381 CTCTGTGCAAAATCGGACCCCAAGATTACATCTCAAGCTGACCAACCAATATTG 4440
4381 CTCTGTGCAAAATCGGACCCCAAGATTACATCTCAAGCTGACCAACCAATATTG 4440
4381 CTCTGTGCAAAATCGGACCCCAAGATTACATCTCAAGCTGACCAACCAATATTG 4440
4441 AAGAGAGAAACCTTCAAAAGACCTTTGAGTCCCGCCCAATGATGTTAGGCGCGG 4500
4441 AAGAGAGAAACCTTCAAAAGACCTTTGAGTCCCGCCCAATGATGTTAGGCGCGG 4500
4501 TCGAGGCGGAGTGTGTCACATGCGGACCAACCACTCCAGGGGAGACAGTAC 4560
4501 TCGAGGCGGAGTGTGTCACATGCGGACCAACCACTCCAGGGGAGACAGTAC 4560
4501 TCGAGGCGGAGTGTGTCACATGCGGACCAACCACTCCAGGGGAGACAGTAC 4560
4561 CGATTAAGTAAATTCGCCAAACATTTCTTAATTTGTAACAAAATTTTAAAAAGAGACA 4620
4561 CGATTAAGTAAATTCGCCAAACATTTCTTAATTTGTAACAAAATTTTAAAAAGAGACA 4620
4621 ATTTTGAATTTGTAACAAATTTTAAACGGGTAATTCCTGAAACATTTTCAAAATTTGT 4680
4621 ATTTTGAATTTGTAACAAATTTTAAACGGGTAATTCCTGAAACATTTTCAAAATTTGT 4680
4621 ATTTTGAATTTGTAACAAATTTTAAACGGGTAATTCCTGAAACATTTTCAAAATTTGT 4680
4681 GATCAAAATTTTAAACGCTCTTCTCAATTTTGAACATTTTAAATTTTAAAAATTA 4740
4681 GATCAAAATTTTAAACGCTCTTCTCAATTTTGAACATTTTAAATTTTAAAAATTA 4740
4681 GATCAAAATTTTAAACGCTCTTCTCAATTTTGAACATTTTAAATTTTAAAAATTA 4740
4741 GTTCAACAATTTGTAATTTTAAATTTGAGGAAACATTTTGAATTTTAAATTTT 4800
4741 GTTCAACAATTTGTAATTTTAAATTTGAGGAAACATTTTGAATTTTAAATTTT 4800
4741 GTTCAACAATTTGTAATTTTAAATTTGAGGAAACATTTTGAATTTTAAATTTT 4800
4801 CGAATTTGGAACATTTTCTTAATTTCTGAACAAAATTTGAAATTTGAAATTTTGA 4860
4801 CGAATTTGGAACATTTTCTTAATTTCTGAACAAAATTTGAAATTTGAAATTTTGA 4860
4801 CGAATTTGGAACATTTTCTTAATTTCTGAACAAAATTTGAAATTTGAAATTTTGA 4860
4861 ATTAATTTTGAAGATTTGGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 4920
4861 ATTAATTTTGAAGATTTGGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 4920
4861 ATTAATTTTGAAGATTTGGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 4920
4921 CTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4980
4921 CTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4980
4921 CTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4980
4981 AAAAGAAATCCGAGAAAAGCACTGGGAATGACATGGAATAAACCAGCGCTCCGCCG 5040
4981 AAAAGAAATCCGAGAAAAGCACTGGGAATGACATGGAATAAACCAGCGCTCCGCCG 5040
4981 AAAAGAAATCCGAGAAAAGCACTGGGAATGACATGGAATAAACCAGCGCTCCGCCG 5040
5041 ACTGTGTAAGCTTAAGTGAAGCGGCGGCAAGCTGCTGCTCATCATCCTCTGTGGA 5100
5041 ACTGTGTAAGCTTAAGTGAAGCGGCGGCAAGCTGCTGCTCATCATCCTCTGTGGA 5100
5041 ACTGTGTAAGCTTAAGTGAAGCGGCGGCAAGCTGCTGCTCATCATCCTCTGTGGA 5100
5101 AACCCGCAATTCGTTGCACTATCGGGGAATGAGCTTTTCAAGAGCTCTGTCTCC 5160
5101 AACCCGCAATTCGTTGCACTATCGGGGAATGAGCTTTTCAAGAGCTCTGTCTCC 5160
5101 AACCCGCAATTCGTTGCACTATCGGGGAATGAGCTTTTCAAGAGCTCTGTCTCC 5160
5161 GGTATAGGTCTATTGCAACCCCTCTCTCACTTGGGCGAGCTTATATCTTTTTC 5220
5161 GGTATAGGTCTATTGCAACCCCTCTCTCACTTGGGCGAGCTTATATCTTTTTC 5220
5161 GGTATAGGTCTATTGCAACCCCTCTCTCACTTGGGCGAGCTTATATCTTTTTC 5220
5221 TTCTTTGCACTGACGTTACTAGCGCAATTTTGAATTTTGAAGCAACCGGTTTGT 5280
5221 TTCTTTGCACTGACGTTACTAGCGCAATTTTGAATTTTGAAGCAACCGGTTTGT 5280

5281 GAAGTCTAAGAACTCAACCATTTTGGAGGCTTCTAGAAGCTATGAATGTTCTT 5340
5281 GAAGTCTAAGAACTCAACCATTTTGGAGGCTTCTAGAAGCTATGAATGTTCTT 5340
5341 TGAACATGATTAATTTGTTGTTTCTTTTCAAAATGCAACATCTTTTCAAAATCAT 5400
5341 TGAACATGATTAATTTGTTGTTTCTTTTCAAAATGCAACATCTTTTCAAAATCAT 5400
5341 TGAACATGATTAATTTGTTGTTTCTTTTCAAAATGCAACATCTTTTCAAAATCAT 5400
5401 GATTTTGTGAACCTGATGATTTTGAATCCGATTTTTCCTAAATCCGTTT 5460
5401 GATTTTGTGAACCTGATGATTTTGAATCCGATTTTTCCTAAATCCGTTT 5460
5401 GATTTTGTGAACCTGATGATTTTGAATCCGATTTTTCCTAAATCCGTTT 5460
5461 GAAAAAACTGTGACCTTTCCGAAATTAATGAACATTTGTTGCAAGATGATCT 5520
5461 GAAAAAACTGTGACCTTTCCGAAATTAATGAACATTTGTTGCAAGATGATCT 5520
5461 GAAAAAACTGTGACCTTTCCGAAATTAATGAACATTTGTTGCAAGATGATCT 5520
5521 TTTCAATGAGCGATTTTCTTAATTAATTCACATATTTTCAATTCATGCTTCC 5580
5521 TTTCAATGAGCGATTTTCTTAATTAATTCACATATTTTCAATTCATGCTTCC 5580
5521 TTTCAATGAGCGATTTTCTTAATTAATTCACATATTTTCAATTCATGCTTCC 5580
5581 TTTTAATGAGTACATCTTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5640
5581 TTTTAATGAGTACATCTTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5640
5581 TTTTAATGAGTACATCTTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5640
5641 TTTTTCCTCAAAAGCAACGTTTAAACGTTGACCGCTGAACCAACCAACCAACCGGGA 5700
5641 TTTTTCCTCAAAAGCAACGTTTAAACGTTGACCGCTGAACCAACCAACCAACCGGGA 5700
5641 TTTTTCCTCAAAAGCAACGTTTAAACGTTGACCGCTGAACCAACCAACCAACCGGGA 5700
5701 AACCATGACCTGCTGAACAGGCGGCTTCTATGATTTGGTGTCTAATACAGCG 5760
5701 AACCATGACCTGCTGAACAGGCGGCTTCTATGATTTGGTGTCTAATACAGCG 5760
5701 AACCATGACCTGCTGAACAGGCGGCTTCTATGATTTGGTGTCTAATACAGCG 5760
5761 CCCCTGACTACTAAGCAAGGAATTTGCAATTTTACCAACCACTACTGATGTTAAAAATG 5820
5761 CCCCTGACTACTAAGCAAGGAATTTGCAATTTTACCAACCACTACTGATGTTAAAAATG 5820
5761 CCCCTGACTACTAAGCAAGGAATTTGCAATTTTACCAACCACTACTGATGTTAAAAATG 5820
5821 AATATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5880
5821 AATATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5880
5821 AATATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5880
5881 ATTTGTTGTTGGGAGAACTTAATCTGAAGAACTTAATCTTAATCTCCGCAATTAATA 5940
5881 ATTTGTTGTTGGGAGAACTTAATCTGAAGAACTTAATCTTAATCTCCGCAATTAATA 5940
5881 ATTTGTTGTTGGGAGAACTTAATCTGAAGAACTTAATCTTAATCTCCGCAATTAATA 5940
5941 TTATGATTCGGGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6000
5941 TTATGATTCGGGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6000
5941 TTATGATTCGGGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6000
6001 ACCCAAGAAAGTCTAGAGTTTGTTCAGAGACCGAAAGATGAGATGCAATTCGA 6060
6001 ACCCAAGAAAGTCTAGAGTTTGTTCAGAGACCGAAAGATGAGATGCAATTCGA 6060
6001 ACCCAAGAAAGTCTAGAGTTTGTTCAGAGACCGAAAGATGAGATGCAATTCGA 6060
6061 ATTAAGTCAATCTTTTCCCAAGATTAAGAAAGATTAAGATTAAGATTAAGATTAAG 6120
6061 ATTAAGTCAATCTTTTCCCAAGATTAAGAAAGATTAAGATTAAGATTAAGATTAAG 6120
6061 ATTAAGTCAATCTTTTCCCAAGATTAAGAAAGATTAAGATTAAGATTAAGATTAAG 6120
6121 GAGATATCATTTCTGATATTAAGTGAATTTGTTTCCGCGCACAGCAACCCCGAGTGG 6180
6121 GAGATATCATTTCTGATATTAAGTGAATTTGTTTCCGCGCACAGCAACCCCGAGTGG 6180
6121 GAGATATCATTTCTGATATTAAGTGAATTTGTTTCCGCGCACAGCAACCCCGAGTGG 6180
6181 AGCCGGAATTTGAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6240
6181 AGCCGGAATTTGAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6240
6181 AGCCGGAATTTGAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6240
6241 TGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6300
6241 TGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6300
6241 TGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6300
6301 GACGATGACGCTGAGGAGGCTGCTGATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 6360
6301 GACGATGACGCTGAGGAGGCTGCTGATGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTG 6360
6301 GACGATGACGCTGAGGAGGCTGCTGATGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTG 6360
6361 TTGATGTGCGGAGGCGAGTGGGACCGGACCGGACCGGCTTAAAGACCGGCCAACGCCG 6420

Db	6361	TTGATGTGCGAGAGAGTGGGAGACGGGACCGACCGCTATTAAAGACCGGCCAACGACCGC	6420
Oy	6421	CGCGGTTCTCAATCCATCCATCCCTTAGCTGATTTGATTGACTAGCTAGTTCAATTCCTG	6480
Db	6421	CGCGGTTCTCAATCCATCCATCCCTTAGCTGATTTGATTGACTAGCTAGTTCAATTCCTG	6480
Oy	6481	CCACACTGTAAGTACTCCTCCGTTTCTGCTGGCATTGGTAACACGACGAAACCGGCA	6540
Db	6481	CCACACTGTAAGTACTCCTCCGTTTCTGCTGGCATTGGTAACACGACGAAACCGGCA	6540
Oy	6541	CGCGAGGCGCGCGCGCGCGCGCCCAACGGACAGCAACGGACACCGCGCGCGCGCA	6600
Db	6541	CGCGAGGCGCGCGCGCGCGCGCCCAACGGACAGCAACGGACACCGCGCGCGCGCA	6600
Oy	6601	CGGCAAGAGCAACGGGGAACGGGGCGGCGCGCGGTGGAGATTTTCGCCCGGGGCA	6660
Db	6601	CGGCAAGAGCAACGGGGAACGGGGCGGCGCGCGGTGGAGATTTTCGCCCGGGGCA	6660
Oy	6661	GGACGGCATCTGTCGCGACGACGGGGGGCAAGAACAGCATCCGGGCGATACGCTAACAGT	6720
Db	6661	GGACGGCATCTGTCGCGACGACGGGGGGCAAGAACAGCATCCGGGCGATACGCTAACAGT	6720
Oy	6721	CAGCGCGACGCTGAGAGAGAGGGGCGCGCGCTGCTGCTGCTGCGCTACCGTGAACC	6780
Db	6721	CAGCGCGACGCTGAGAGAGAGGGGCGCGCGCTGCTGCTGCTGCGCTACCGTGAACC	6780
Oy	6781	GTCCGCTGTTCCCGGCGCTTCGCGACGGGCGGTGAGAGGACGGCGCTCCCGCGCGCGGCT	6840
Db	6781	GTCCGCTGTTCCCGGCGCTTCGCGACGGGCGGTGAGAGGACGGCGCTCCCGCGCGCGGCT	6840
Oy	6841	GGCGACCGGCGAGTTCAACTGCTACCGCGCGCGCGCTCCCGCGCGCACAGGTA	6900
Db	6841	GGCGACCGGCGAGTTCAACTGCTACCGCGCGCGCGCTCCCGCGCGCACAGGTA	6900
Oy	6901	ACATTTCACGCTTCAACCGTAAATGTAATGGGTAGACATGATGCGCGGTTTACTTAAGTGC	6960
Db	6901	ACATTTCACGCTTCAACCGTAAATGTAATGGGTAGACATGATGCGCGGTTTACTTAAGTGC	6960
Oy	6961	CGCGCGCTGTTCTTCCCGGTCGCTCAAAATTTTAACTTTAACTTAAGTACTTATAA	7020
Db	6961	CGCGCGCTGTTCTTCCCGGTCGCTCAAAATTTTAACTTTAACTTTAACTTATAA	7020
Oy	7021	ACAAACAGCGCGCTAGACAGACCTTGTCAACAGGCGTGCCTTACAGCTATCGGCGAC	7080
Db	7021	ACAAACAGCGCGCTAGACAGACCTTGTCAACAGGCGTGCCTTACAGCTATCGGCGAC	7080
Oy	7081	GACGCTTCTCAACCGCGCGGCGGAATCGAGGCAATCGAAGTCATAATCCGCGTCTGGCC	7140
Db	7081	GACGCTTCTCAACCGCGCGGCGGAATCGAGGCAATCGAAGTCATAATCCGCGTCTGGCC	7140
Oy	7141	CAGAATCGCGGCGGCAATATCTGCTTCCCGGCGAGGCTATCCAAATTAAGAGCGCA	7200
Db	7141	CAGAATCGCGGCGGCAATATCTGCTTCCCGGCGAGGCTATCCAAATTAAGAGCGCA	7200
Oy	7201	GCGGCAATTCACAAAGCTGGAAGTCCGGCACTTTCGACCTCATCCCGGACAAAGGGTGGAG	7260
Db	7201	GCGGCAATTCACAAAGCTGGAAGTCCGGCACTTTCGACCTCATCCCGGACAAAGGGTGGAG	7260
Oy	7261	ATCGACATCGACTGCTGGAATTCATCGCGACAAAGAACACACACGCGAGTGTATGATA	7320
Db	7261	ATCGACATCGACTGCTGGAATTCATCGCGACAAAGAACACACACGCGAGTGTATGATA	7320
Oy	7321	AAACCCAAACATCCGTCGCGAGAGGTTTACTTCACACATCTGCGGCAAGGTTTGAT	7380
Db	7321	AAACCCAAACATCCGTCGCGAGAGGTTTACTTCACACATCTGCGGCAAGGTTTGAT	7380
Oy	7381	CCATGCACTCTCTGCTGTTGATCGACCGGCTGTTTGAACATAGTATAGATTGCGT	7440
Db	7381	CCATGCACTCTCTGCTGTTGATCGACCGGCTGTTTGAACATAGTATAGATTGCGT	7440
Oy	7441	TTGCTAATCGTGTGCTGATGATGCTGTTTGGTATCAGGTCGCGAGGTGGCAAGAAC	7500
Db	7441	TTGCTAATCGTGTGCTGATGATGCTGTTTGGTATCAGGTCGCGAGGTGGCAAGAAC	7500

D	7441	TTGGCTAATCGTGTGCTATGATGCTGTTTGGTTATCAAGTGGCGGAAGTGGCAAGAAC	7500
Q	7501	TCGGAAATTTGGTGAATGCTGCTGACGAGGGTTTACGGCAAACTGCTTCTGGGCAAGCCGCCCT	7560
D	7501	TCGGAAATTTGGTGAATGCTGCTGACGAGGGTTTACGGCAAACTGCTTCTGGGCAAGCCGCCCT	7560
Q	7561	TTATCCCGAATGGGCGCTCTTTGGGACATATGCGCCGGCTCTTGTCATTGGATCTCTGATCCA	7620
D	7561	TTATCCCGAATGGGCGCTCTTTGGGACATATGCGCCGGCTCTTGTCATTGGATCTCTGATCCA	7620
Q	7621	AGTCGTGATAGTGCCTGGATGAGCGACTTGGATGGGTGGCGGTATGACGACCCCAAGA	7680
D	7621	AGTCGTGATAGTGCCTGGATGAGCGACTTGGATGGGTGGCGGTATGACGACCCCAAGA	7680
Q	7661	TTTTAGAGAAACCTAAGGTACTCTTAGTCCCTATCAATCTCTCATATGTACTATGGGG	7740
D	7661	TTTTAGAGAAACCTAAGGTACTCTTAGTCCCTATCAATCTCTCATATGTACTATGGGG	7740
Q	7741	GATTAGTATTTTGTGCTAAATTTGTACTGCCCTTGTATTACAGATCTCTACGTCTATTAC	7800
D	7741	GATTAGTATTTTGTGCTAAATTTGTACTGCCCTTGTATTACAGATCTCTACGTCTATTAC	7800
Q	7801	GAATTTACTTATATGCTCTACAGGACCCAGCAACTTGTTCAGGTATGCTTTGGTCTT	7860
D	7801	GAATTTACTTATATGCTCTACAGGACCCAGCAACTTGTTCAGGTATGCTTTGGTCTT	7860
Q	7861	GCCCTATTTTGTCTATGTCCTCTGTGTGTCATGTCAAAATGACCGGCTTCAAGTTAGTAT	7920
D	7861	GCCCTATTTTGTCTATGTCCTCTGTGTGTCATGTCAAAATGACCGGCTTCAAGTTAGTAT	7920
Q	7921	AGAGTTTGTGTAAGTGAATGTGAAGTCCACATGATGGAAGAAAGATACCTATT	7980
D	7921	AGAGTTTGTGTAAGTGAATGTGAAGTCCACATGATGGAAGAAAGATACCTATT	7980
Q	7961	TTTAGTCAATCCCTGTTTGTGTTGTTGATCCATAAATTAATTAACAACAAGCCAGAAC	8040
D	7961	TTTAGTCAATCCCTGTTTGTGTTGTTGATCCATAAATTAATTAACAACAAGCCAGAAC	8040
Q	8041	AACATTGAATAGAACTATTTTCTTAGAAATATATGATGTATTTTGAGCATGCCATAT	8100
D	8041	AACATTGAATAGAACTATTTTCTTAGAAATATATGATGTATTTTGAGCATGCCATAT	8100
Q	8101	TCTTTTGCATCAAGTATGCAATATATTTAAACCTGCAATGTACTACAGATATACCATGT	8160
D	8101	TCTTTTGCATCAAGTATGCAATATATTTAAACCTGCAATGTACTACAGATATACCATGT	8160
Q	8161	GTTAAGAATTTCTTACCTACACACCTGTCTGCACTTCAATTTTGATCTCTTGA	8220
D	8161	GTTAAGAATTTCTTACCTACACACCTGTCTGCACTTCAATTTTGATCTCTTGA	8220
Q	8221	CATTATGTTCTCTTATGATTCACACAACTTAATATGATTTTGTGCTATCAAAATGT	8280
D	8221	CATTATGTTCTCTTATGATTCACACAACTTAATATGATTTTGTGCTATCAAAATGT	8280
Q	8281	TTAGGAAGCTCTCTAAATTTCTTGGAACACAAAGCAGATTTCTTTAAGAGATTAAT	8340
D	8281	TTAGGAAGCTCTCTAAATTTCTTGGAACACAAAGCAGATTTCTTTAAGAGATTAAT	8340
Q	8341	TGCTCTACTAAGGAATCATCAGATATGTTATAGGAAATTAAGAAAACAATATAT	8400
D	8341	TGCTCTACTAAGGAATCATCAGATATGTTATAGGAAATTAAGAAAACAATATAT	8400
Q	8401	TACGTGTCTCACAAGCCAGAAAGATGATGTTGTATGGTAAGCTTAAGATATGACTTA	8460
D	8401	TACGTGTCTCACAAGCCAGAAAGATGATGTTGTATGGTAAGCTTAAGATATGACTTA	8460
Q	8461	CTTTTAAAGTTATCTGGGATCTCAAGTCATCCAAACAATCAATCAAAATATTAAT	8520
D	8461	CTTTTAAAGTTATCTGGGATCTCAAGTCATCCAAACAATCAATCAAAATATTAAT	8520
Q	8521	TATGTTTGTCTATGATCTTTTGAAGATGATCTTTGAAGATTAATGAAGAGATGG	8580
D	8521	TATGTTTGTCTATGATCTTTTGAAGATGATCTTTGAAGATTAATGAAGAGATGG	8580

8581 AAATATTTAGAGCTAATCTTCTGATATCATTTGTGCAATTTTGTATTAATCTGAA 8640
 8582 AATATTTTGAAGACTAATCTTCTGATATCATTTGTGCAATTTTGTATTAATCTGAA 8640
 8641 ATGGTAACACTCAATCATATTTACAAAAGTTTCTCCCATTTTGTAGATGACTT 8700
 8642 ATGGTAACACTCAATCATATTTACAAAAGTTTCTCCCATTTTGTAGATGACTT 8700
 8701 CCTTCTAATCAAGATTAATCTTCAATGTAAGAGTCAATCAATCAATCTTCTT 8760
 8702 CCTTCTAATCAAGATTAATCTTCAATGTAAGAGTCAATCAATCAATCTTCTT 8760
 8703 CCTTCTAATCAAGATTAATCTTCAATGTAAGAGTCAATCAATCAATCTTCTT 8760
 8761 GGAGAGATTCATGACGACATTAATTTTGTGCAAGCTCGCAAGAGAAATCTGTAT 8820
 8762 GGAGAGATTCATGACGACATTAATTTTGTGCAAGCTCGCAAGAGAAATCTGTAT 8820
 8821 TTTATGTCAGAGTAAATGTAATGTCCTTTTAAAGAAATCAATATGAAATATAT 8880
 8822 TTTATGTCAGAGTAAATGTAATGTCCTTTTAAAGAAATCAATATGAAATATAT 8880
 8881 ATCTCTGTAATTAATTAATTAATCTTCTCAATCTTCAATTAATTAATTAATTAAT 8940
 8882 ATCTCTGTAATTAATTAATTAATCTTCTCAATCTTCAATTAATTAATTAATTAAT 8940
 8941 ACAATTTTAACTAGATCTAGTACAAAGTTGAAACAGTTATTTGGGACAGAGGAGT 9000
 8942 ACAATTTTAACTAGATCTAGTACAAAGTTGAAACAGTTATTTGGGACAGAGGAGT 9000
 9001 TATATATGTTGAGAACATTAAGTTTGTGATGATATATGCTTCTTAAATGTGAA 9060
 9002 TATATATGTTGAGAACATTAAGTTTGTGATGATATATGCTTCTTAAATGTGAA 9060
 9061 CATGTTCTTAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 9120
 9062 CATGTTCTTAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 9120
 9121 ACATTAATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9180
 9122 ACATTAATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9180
 9181 TTTTATGCAAGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 9240
 9182 TTTTATGCAAGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 9240
 9241 TCGTAAATGTTGTTGCTTCAAAAATTAATTAATCAAGCAATGTTCCAAATGAT 9300
 9242 TCGTAAATGTTGTTGCTTCAAAAATTAATTAATCAAGCAATGTTCCAAATGAT 9300
 9301 ATTAAATTTTGAAGTGTGTTTCAACCAATTAATTAATTAATTAATTAATTAATTAAT 9360
 9302 ATTAAATTTTGAAGTGTGTTTCAACCAATTAATTAATTAATTAATTAATTAATTAAT 9360
 9361 GATATATGTTGATCTTAACCAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 9420
 9362 GATATATGTTGATCTTAACCAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 9420
 9421 TCTTGAATGAAAAATGGGCTGATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9480
 9422 TCTTGAATGAAAAATGGGCTGATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9480
 9481 TGGACTCGAAGGCTCAATCTTCTGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 9540
 9482 TGGACTCGAAGGCTCAATCTTCTGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 9540
 9541 TTTGATGTTGTAACACCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 9600
 9542 TTTGATGTTGTAACACCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 9600
 9601 ATCCCATTTATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9660
 9602 ATCCCATTTATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9660

9661 CATGATCTACTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 9720
 9662 CATGATCTACTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 9720
 9721 GGTGTTGTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9780
 9722 GGTGTTGTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9780
 9781 CTATATCTTTCAACCAATGAAACCTGAGTCCCTTGAAGTGAAGTGAAGTGAAGTGAAG 9840
 9782 CTATATCTTTCAACCAATGAAACCTGAGTCCCTTGAAGTGAAGTGAAGTGAAGTGAAG 9840
 9841 GAATATGCTATGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9900
 9842 GAATATGCTATGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9900
 9901 CATGCAAGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9960
 9902 CATGCAAGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9960
 9961 ACATTTCTGAAGTAAATGTCAGTTGGTGAAGGACCTTAATTTTCTATTTGGG 10020
 9962 ACATTTCTGAAGTAAATGTCAGTTGGTGAAGGACCTTAATTTTCTATTTGGG 10020
 10021 GGGGGGGGGTGAATAGCGGTTTAAACCAATGTAATTTGAGAAATCTTAATGTTGA 10080
 10022 GGGGGGGGGTGAATAGCGGTTTAAACCAATGTAATTTGAGAAATCTTAATGTTGA 10080
 10081 ATTAACTAGTGAATTTTTCCTTCAATTAAGGTTGCTTTATGATCTCAATGTTACA 10140
 10082 ATTAACTAGTGAATTTTTCCTTCAATTAAGGTTGCTTTATGATCTCAATGTTACA 10140
 10141 TCAAGGATCAATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 10200
 10142 TCAAGGATCAATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 10200
 10201 CCAAC 10260
 10202 CCAAC 10260
 10261 GTCAATCAAGTCAAACTAATGCTAGCGGAGGAGAAATGAAGAAATGAAGAAATGA 10320
 10262 GTCAATCAAGTCAAACTAATGCTAGCGGAGGAGGAGAAATGAAGAAATGAAGAAATGA 10320
 10321 AAAACCGTGAACATCTAGTCAATGATGATGATGATGATGATGATGATGATGATGATG 10380
 10322 AAAACCGTGAACATCTAGTCAATGATGATGATGATGATGATGATGATGATGATGATG 10380
 10381 CTGCGAAGAAATGTTCTATTAATTAATGAGGCTTTCGAGAAAGAAAGAGAGAGAGAG 10440
 10382 CTGCGAAGAAATGTTCTATTAATTAATGAGGCTTTCGAGAAAGAAAGAGAGAGAGAG 10440
 10441 CATCATGGAATCAACCACTAAGGCTTAATGAGGCTTTCGAGAAAGAGAGAGAGAGAG 10500
 10442 CATCATGGAATCAACCACTAAGGCTTAATGAGGCTTTCGAGAAAGAGAGAGAGAGAG 10500
 10501 AGCAAACTCGAGTAAATGCTTTATTAAGGCTTAATGAGGCTTTCGAGAAAGAGAGAGAG 10560
 10502 AGCAAACTCGAGTAAATGCTTTATTAAGGCTTAATGAGGCTTTCGAGAAAGAGAGAGAG 10560
 10561 TTAATGCAATTAATTAACCAAGCTTGAATTTTATTCGAAGCTCGAGAGGAGTACTTGA 10620
 10562 TTAATGCAATTAATTAACCAAGCTTGAATTTTATTCGAAGCTCGAGAGGAGTACTTGA 10620
 10621 AGTACATCAATTAATTAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10680
 10622 AGTACATCAATTAATTAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10680
 10681 GTCCATATGCGAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10740
 10682 GTCCATATGCGAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10740
 10741 GCTGTGCGCCACCACTGACGAAGTGTGATGAGGAGAGAAAGATATGGCTCCACCAAC 10800

OY 1717 CGACATGCACTCGTGAATCATGCGGACAGAACACCCGGCATGTCATCATATAA 1776
 DB 960 CGACATGCACTCGTGAATCATGCGGACAGAACACCCGGCATGTCATCATATAA 1019
 OY 1777 CCCCACAAACCCGTCGCGGACGGCTTACTCTACACCATCTGTCCAGTTTCACATCC 1836
 DB 1020 CCCCACAAACCCGTCGCGGACGGCTTACTCTACACCATCTGTCCA----- 1066
 OY 1837 TTTCGCTTGTGAATATGATTCAGTTCAGTCCACTGCTGATTTCTTTGCCAATGCC 1896
 DB 1067 ----- 1066
 OY 1897 ATACTGACGATGTTGCTCAATTAGTTCGCGGAGGTCGGAAGAGCTCGAATATTGAT 1956
 DB 1067 -----AGTCCGAGAGGTGGGAGGAGGCTCGAATATTGAT 1103
 OY 1957 GATTGCTGACGAGATATACGGCAAGCTGTTCTGGGACAGCCCGCTTATCCCAATGGG 2016
 DB 1104 GATTGCTGACGAGATATACGGCAAGCTGTTCTGGGACAGCCCGCTTATCCCAATGGG 1163
 OY 2017 AGTGTTCGGGACATACACCCCTGCTGTCTCTAGAGGTTCTCTTCAAGTCATGATAGT 2076
 DB 1164 AGTGTTCGGGACATACACCCCTGCTGTCTCTAGAGGTTCTCTTCAAGTCATGATAGT 1223
 OY 2077 GCCTGATGCGCGCTTGGATGGGTAGCGGCTGACACCCAGAAAGATCTTACAGAGAAC 2136
 DB 1224 GCCTGATGCGCGCTTGGATGGGTAGCGGCTGACACCCAGAAAGATCTTACAGAGAAC 1283
 OY 2137 TAAGTACTTAATC 2151
 DB 1284 TAAGTACTTAATC 1298

RESULT 3

AAV48147 standard; cDNA to mRNA; 1660 BP.

AAV48147;

27-AUG-2003 (revised)
27-OCT-1998 (first entry)

DE Nicotianamine aminotransferase 49564.15 molecular weight protein, gene.

KM ds; nicotianamine aminotransferase; plant; iron absorption;
iron deficiency chlorosis.

OS Poaceae.

XX Location/Qualifiers

FT Key 62.1447
FT CDS /tag= a
FT /product= "Nicotianamine aminotransferase"

PN EP860499-A2.

PD 26-AUG-1998.

PF 19-FEB-1998, 98EP-00102891.

PR 21-FEB-1997, 97JP-00037499.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX Mori S, Nakanishi H, Takahashi M,

XX WPI, 1998-439341/38.

DR P-PsDB; AAM61642.

PT New nicotianamine aminotransferase protein and DNA - useful for enhancing
iron absorption of plant cells.

PS Claim 4; Page 12-13; 17pp; English.

CC The nicotianamine aminotransferase can be used in a plasmid to transform
 CC plant cells to produce cells with enhanced iron absorption, and it is
 CC implied (though not stated) that plants with improved resistance to iron
 CC deficiency chlorosis in calcareous soils can be regenerated from the
 CC transformed cells. The gene fragment can be used to detect, amplify
 CC and/or isolate nicotianamine aminotransferase genes. (Updated on 27-AUG-
 CC 2003 to correct OS field.)

XX Sequence 1660 BP; 423 A; 442 C; 430 G; 365 T; 0 U; 0 Other;

Query Match 6.8%; Score 744.6; DB 2; Length 1660;

Best Local Similarity 80.6%; Pred. No. 1e-115;

Matches 1005; Conservative 0; Mismatches 4; Indels 238; Gaps 2;

OY 6457 ATTGACTGAGTATGTTCAATTCCTTCCCACTGCTAGTACTCTCTCTGTTCTGTCGGC 6516
 DB 1 ATTGACTGAGTATGTTCAATTCCTTCCCACTGCTAGTACTCTCTCTGTTCTGTCGGC 60
 OY 6517 AATGTTACACGAGCAAGGCGACAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6576
 DB 61 AATGTTACACGAGCAAGGCGACAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 OY 6577 CAACGCGCACG 6636
 DB 121 CAACGCGCACG 180
 OY 6637 GGAATGGAATTTCCCG 6696
 DB 181 GGAATGGAATTTCCCG 240
 OY 6697 CATCCGCGCGATGAGTAAAGATCAAGATCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6756
 DB 241 CATCCGCGCGATGAGTAAAGATCAAGATCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 OY 6757 GCTGCGCGTGGCCACGAGTACCCGTCGTTCCCGCGCTTCCGACGCGCGCGTGAAGC 6816
 DB 301 GCTGCGCGTGGCCACGAGTACCCGTCGTTCCCGCGCTTCCGACGCGCGCGTGAAGC 360
 OY 6817 CGAGAGGCGCGTGGCG 6876
 DB 361 CGAGAGGCGCGTGGCG 420
 OY 6877 CGGCTCCCG 6936
 DB 421 CGGCTCCCG 439
 OY 6937 GCATGCGCGGTTTACTTAAGTACGTCGCGCGCGCTGTTCTTCCCGGTCGTTCAAAATTTTA 6996
 DB 440 ----- 439
 OY 6997 ACCTTCTAATAGTACTTATATAAAACAAACGCGCGGTAGAGAGCACTTGTCAAGGCG 7056
 DB 440 -----AGCGCGTAGAGAGCACTTGTCAAGGCG 469
 OY 7057 GTGCGCTTAAAGTATGCGCGCGAGAGAGTCTTCTCAACCGCGCGGAGAACTCAAGCGATC 7116
 DB 470 GTGCGCTTAAAGTATGCGCGCGAGAGAGTCTTCTTCTCAACCGCGCGGAGAACTCAAGCGATC 529
 OY 7117 GAAATCATTAATCCCGGTGCTGGCCAGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7176
 DB 530 GAAATCATTAATCCCGGTGCTGGCCAGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 589
 OY 7177 GGCATTCGAATTTACGAGCG 7236
 DB 590 GGCATTCGAATTTACGAGCG 649
 OY 7237 CTGATCCCGCGCAAGGCGGTTGGAGATGCAATGCACTGCTGGAATTCATGCGCGCAAG 7296
 DB 650 CTGATCCCGCGCAAGGCGGTTGGAGATGCAATGCACTGCTGGAATTCATGCGCGCAAG 709
 OY 7297 AACACACCGGAGTGTATCATTAACCAACCAATCGTGGCGAGCGTTTACTCTTAC 7356

```

Db 710 AACACACCGCATGATCATCAATAACCAACCAATCCGTGCGGAGCGTTTACTCTAC 769
Qy 7357 GACCATTCGGCCCAAGTTTGATCCATCCCTCTGCTGTTGATGACCGCTCTCT 7416
Db 770 GACCATTCGGCCCAAGTTTGATCCATCCCTCTGCTGTTGATGACCGCTCTCT 782
Qy 7417 TTGAACATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7476
Db 783 ----- 782
Qy 7477 AGGTGCGGAGGTGCGCAAGAGCTCGAATAATTGCTGATCGCTGACGAGTTTACGCGCA 7536
Db 783 AGGTGCGGAGGTGCGCAAGAGCTCGAATAATTGCTGATCGCTGACGAGTTTACGCGCA 842
Qy 7537 AACTGCTTTCGGGACGCGCCCGCTTATCCGATGCGGCTCTTTGGGCAATTGCCCCG 7596
Db 843 AACTGCTTTCGGGACGCGCCCGCTTATCCGATGCGGCTCTTTGGGCAATTGCCCCG 902
Qy 7597 TCTTGTCCATTTGATCTCTGTCGAAGTCGTGATAGTGTGCTGATGCGACTTGGATGG 7656
Db 903 TCTTGTCCATTTGATCTCTGTCGAAGTCGTGATAGTGTGCTGATGCGACTTGGATGG 962
Qy 7657 TGGCGGTGTACGACCCCAAAAGATTTTGAAGAAACTAGGTACT 7703
Db 963 TGGCGGTGTACGACCCCAAAAGATTTTGAAGAAACTAGGTACT 1009

```

RESULT 4
ADKS2801
ID ADKS2801 standard; DNA; 568 BP.

ADKS2801;

06-MAY-2004 (first entry)

Plant DNA sequence which confers altered metabolic characteristic #184.

altered metabolic characteristic; plant; acid metabolism;
alcohol metabolism; fatty acid metabolism;
branched fatty acid metabolism; alkaloid metabolism;
amino acid metabolism; ester metabolism; glyceride metabolism;
phenolic metabolism; carbohydrate metabolism; steroid metabolism;
terpene metabolism; isoprenoid metabolism; alkene metabolism;
alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.

Oryza sativa.

WO2003020936-A1.

13-MAR-2003.

30-AUG-2002; 2002WO-US027884.

31-AUG-2001; 2001US-0316471P.

(DOMC) DOM CHEM CO.

(DOMC) DOM AGRSCIENCES LLC.

Weglarz T, Gachotte D, Blakelee B, McCreary DA, Pell RJ,

Orlando JVB, Crooley R, Reddy AS, Shukla V, Larrina I, Miller BA,

WPI; 2003-313091/30.

Novel genes that confer altered metabolic characteristics in Nicotiana
benhamiana plants, useful for altering the levels of metabolites e.g.
acids, fatty acids, amino acids, carbohydrates, hydrocarbons and steroids.

Claim 1; SEQ ID NO 184; 2576bp; English.
The invention comprises DNA sequences which confer an altered metabolic
characteristic when they are expressed in a plant. The DNA sequences of

CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered steroid, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.

Sequence 568 BP; 113 A; 201 C; 171 G; 83 T; 0 U; 0 Other;

Query Match 1.7%; Score 184.4; DB 10; Length 568;
Best Local Similarity 78.9%; Pred. No. 26-21;
Matches 232; Conservative 0; Mismatches 60; Indels 2; Gaps 1;

```

Qy 1008 CGCGCGCGCGCGCAGAGAGAGAGAGC--GGTGAAGTGAATTTCGCGGTGCCAAGA 1065
Db 64 CGCGCGCGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
Qy 1066 CGCGGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1125
Db 124 CGCGCGCGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 183
Qy 1126 CGCGAGCGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
Db 184 CGCGAGCGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
Qy 1186 CGTGTTCGCGCGCTTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1245
Db 244 CGTGTTCGCGCGCTTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 303
Qy 1246 CACCGCGCAGTTCACTGCTACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1299
Db 304 CACCGCGCAGTTCACTGCTACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 357

```

RESULT 5

ADA70556
ID ADA70556 standard; DNA; 1233 BP.

ADA70556;

20-NOV-2003 (first entry)

Rice gene, SEQ ID 3879.

Plant; bacterial infection; fungal infection; viral infection; rice;
gene; ds.

Oryza sativa.

WO200300898-A1.

03-JAN-2003.

22-JUN-2001; 2001WO-IB001105.

22-JUN-2001; 2001WO-IB001105.

(SYGN) SYNGENTA PARTICIPATIONS AG.

Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,

Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;

WPI; 2003-175290/17.

Identifying at least one gene involved in plant resistance or response to
pathogenic infection for conferring resistance or tolerance to a plant to
bacterial, fungal or viral infection by determining or detecting plant
gene expression.

XX
PS Claim 6; SEQ ID NO 3879; 899pp; English.

The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to the expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to illustrate the invention.

Sequence 1233 BP; 221 A; 397 C; 412 G; 201 T; 0 U; 2 Other;

Query Match 1.3%; Score 147; DB 8; Length 1233;
Best Local Similarity 67.0%;
Matches 223; Conservative 1; Mismatches 106; Indels 3; Gaps 1;

OY	1480	CAGCGCGGTGGCAGAGCACTGTGGCAAGGGGTGGCGTAAATGCTATGGCGCAACGACT	1539
Db	294	CAGCGCGGTGGCGGCGCACTGTGTGGCGGAGGTCCCCCTTAAGCCGCTCTGGCGGCGAGCT	353
OY	1540	CTTCTCAACCGCGCGGAGACCAGAGCATGAGGTCAATATCCGGTGTGGGCCAGAC	1599
Db	354	GGTGTCTACCGCGCGGTCTCAACACAGCGGTGACAGATCATGTATGTCTGGTCTGC---GTC	410
OY	1600	CGCGCGCGCCAAATTTGTGTCTCCCGAGGCGAGGCTAACCAATCTACGAGGGCGCGCGCG	1659
Db	411	GCCGGGCGCCAACTGTGTGTCTCTCCGGGCGCGGCTTACCGCGCTGTACGCGGTCCGCGCGCG	470
OY	1660	GTTCAACAGGCTGAGGTTCGGCATTTTGCACCTCATTCCCGCAACAAGGGTGGAGATCGA	1719
Db	471	CCTGAGCGGCGCTGTGAGTTCGCGCATTTGGAACCTCTCTCCCGCAGCGAGTGGAGGTGTGA	530
OY	1720	CATGCACTCGCTGGAATCATGCGCGCAACAAGAACCGCGCATGTGATCATTAACCC	1779
Db	531	CCTGCGCGGCGTGTGAGGCGCTCTGCCGAGCCCAAAACCGTGTCCATGATTCGAGAGACGGC	590
OY	1780	CAACAACCGGTGGCGGAGCGTTTACTCTTAACA	1812
Db	591	GAGGAAGCTGGGAGATTAATGTGTATAGGACACA	623

RESULT 6

ID ABZ14064 standard; DNA; 1338 BP.

AC ABZ14064 ;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1869.

KW *Arabidopsis thaliana*; plant; gene; stress; transgenic; ds.

OS *Arabidopsis thaliana*.

PN WO200216655-A2.

PD 28-FEB-2002

PF 24-AUG-2001; 2001WO-US026685.

PR 24-AUG-2000; 2000US-0227866P.

PR 22-JUN-2001; 2001US-0300111P.

PA (SCRI) SCRIPPS RES INST.

2X

XX WPI; 2002-304127/34.
DR

PT Identifying a stress condition to which a plant cell has been exposed and PT
PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 1869; 577pp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an *Arabidopsis thaliana* stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office

Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 U; 0 Other;

Query Match	1.3%	Score 142.2;	DB 6;	Length 1338;
Best Local Similarly	63.1%	Pred. No. 3.2e-14;		
Matches 219; Conservative	0;	Mismatches 128;	Indels 0;	Gaps 0;

QY	1481	AGGCGCGGTGGCAGAGCACTGTGGCAGGGCGGTGGCCGATCAATGATCGGCGCGAGCGATC	1540
Db	292	AGGGCGGTGGCTGTAATATTATTTAAACGAGAACTTCGACGAACTGTAAGGCCGAGATGTG	351
OY	1541	TTCTCTAACCGCGCGGAGACCCGAGCGCATGAGGTATATATCCCGTGTCTG3CCAGACC	1600
Db	352	TATATCAACCGAGGATGTATCAACGACCATAGAGATCGATAGATATCTCTTCCCGGAAT	411
OY	1601	GC CGCGCGCAATTTCTGTCTCCCGAGGCGAGCTACCCAACTACGAGCGCGCGCGCG	1660
Db	412	CCATCCCGCAATTTCTACTTTCAGAGCGGGATATCTCTACCTACGATCTGTGTCTTC	471
OY	1661	TTCAACAGGCTGAGGTCCGGCTTTTGACTCATATCCCGACAAAGGGTGGAGATGCAC	1720
Db	472	TATAGCGGCTTCGAGATTCGCAAAATCGATTTCTCCCGAGATGATTTGGAAATCAAT	531
OY	1721	ATCGACTCGCTGGAATCATCGCCGACAAAGAACACCGCATGTGATCATATAAACCC	1780
Db	532	CTCATGTGCTTCAGAGCGGCTGCGGATGAGAAATCCGTGCAATGTGTAATCATCAACCC	591
OY	1781	AACAACCGGTGGCGAGCGTTTACTCTCTACGACCATGTCTCCAAAGT	1827
Db	592	AACATCATGTGAAACGTCTACACCTACACCATCTCCAAACAGGT	638

RESULT 7

ID ADG87698 standard; cDNA; 1338 BP.

AC ADG87698

DT 22-APR-2004 (first entry)

A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #140.

KW Pathogen infection-related gene; plant; *Peronospora parasitica*;

KW oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.

OS *Arabidopsis thaliana*.

PN . WO200222675-A2.

PD 21-MAR-2002

PF 14-SEP-2001; 2001WO-US028506.

XX 15-SEP-2000; 2000US-0232778P.
 PR 22-JUN-2001; 2001US-0300183P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA (UYN-) UNIV NORTH CAROLINA.
 PA (GLAZ/) GLAZEBROOK J.
 PA (WANG/) WANG X.
 PA (DANG/) DANG J L.
 PA (EUG/) EUGEN T.
 PA (ZHUT/) ZHU T.
 XX
 PI Glazebrook J, Wang X, Dang J, Eugen T, Zhu T;
 XX
 DR MPI; 2002-292409/33.
 XX
 PT Novel isolated polynucleotide, useful for conveying pathogen resistance
 XX to plants, and for identifying plants infected with a pathogen.
 PS
 XX Claim 3: SEQ ID NO 140; 605bp; English.
 CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
 CC ADG87557) whose expression is altered in response to pathogen infection,
 CC and to homologues of these genes from other plants or fungi, especially
 CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
 CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
 CC expression of genes of the invention was upregulated or downregulated in
 CC Arabidopsis plants infected with the comycete *Peronospora parasitica*,
 CC indicating that they play a role in defence mechanisms. The genes of the
 CC invention are regulated by RPP7 or RPP8 which act via unconventional
 CC signalling cascades, or by the RPP4-dependent pathway. The invention also
 CC relates to polypeptides encoded by the pathogen infection-related genes;
 CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
 CC / expression cassettes, host cells and pathogen-resistant transgenic
 CC plants and their progeny comprising a polynucleotide of the invention;
 CC and a method of identifying a plant cell infected with a pathogen. The
 CC polynucleotide sequences and methods of the invention are useful for
 CC identifying plants infected with a pathogen, and for conferring
 CC resistance to pathogens such as comycetes, fungi, bacteria, viruses,
 CC nematodes and insects (e.g., aphids). The present sequence represents an
 CC Arabidopsis thaliana gene whose expression is altered in response to
 CC *Peronospora parasitica* infection. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 U; 0 Other;
 Query Match 1.3%; Score 142.2; DB 6; Length 1338;
 Best Local Similarity 63.1%; Pred. No. 3.2e-14;
 Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
 Oy 1481 AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 1540
 Db AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 1540
 Oy 292 AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 351
 Db AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 351
 Oy 1541 TTCCTACCGCGCGCGGAGACCGAGCGATCGAGTCATATCCCGGTGTCGCCAGAC 1600
 Db TTCCTACCGCGCGCGGAGACCGAGCGATCGAGTCATATCCCGGTGTCGCCAGAC 1600
 Oy 352 TATATACCGCGAGATGTAACCAAGCCATAGAGATCGATGATTTCTTTCGCGAAT 411
 Db TATATACCGCGAGATGTAACCAAGCCATAGAGATCGATGATTTCTTTCGCGAAT 411
 Oy 1601 GCCCGCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 1660
 Db GCCCGCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 1660
 Oy 412 CCATCCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 471
 Db CCATCCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 471
 Oy 1661 TTCAACAGGCTGAGAGTCGCGCATTTGCACTCATTCCTCCGACAGAGGGTGGAGATGAC 1720
 Db TTCAACAGGCTGAGAGTCGCGCATTTGCACTCATTCCTCCGACAGAGGGTGGAGATGAC 1720
 Oy 472 TATACCGCGCTCGAATTCGCAAAATACGATCTTCTCCCGAGAGATGGGAATCAAT 531
 Db TATACCGCGCTCGAATTCGCAAAATACGATCTTCTCCCGAGAGATGGGAATCAAT 531
 Oy 1721 ATCGAGCTGAGAGTCATGCGCGAGCAAGACACACCGCATGTCATCAATACCC 1780
 Db ATCGAGCTGAGAGTCATGCGCGAGCAAGACACACCGCATGTCATCAATACCC 1780
 Oy 532 CTCGATGCGCTCGAGCGCGCGATGAGATACCGTCGCAATGATATCAATCAACCC 591
 Db CTCGATGCGCTCGAGCGCGCGATGAGATACCGTCGCAATGATATCAATCAACCC 591
 Oy 1781 AACCAACCGCTGCGAGCGGCTTACTCTCAACGACATCTGTCCAAGT 1827

Db 592 AACCATCCATGTGGAACGCTTACCTTACGACCATCTCAAGAGT 638
 RESULT 8
 ADA68509
 ID ADA68509 standard; DNA; 1338 BP.
 XX
 AC ADA68509;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Arabidopsis thaliana gene, SEQ ID 584.
 XX
 KM Plant; bacterial infection; fungal infection; viral infection; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN MO200300898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-1B001105.
 XX
 PR 22-JUN-2001; 2001WO-1B001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Qian S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
 XX
 DR MPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 PS Claim 6; SEQ ID NO 584; 899bp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 U; 0 Other;
 Query Match 1.3%; Score 142.2; DB 8; Length 1338;
 Best Local Similarity 63.1%; Pred. No. 3.2e-14;
 Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
 Oy 1481 AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 1540
 Db AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 1540
 Oy 292 AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 351
 Db AGCGCGTGGAGACACCTGTGCGAGGCGGTGCGTACATGCTATCGCGCGACGACGTC 351
 Oy 1541 TTCCTACCGCGCGGAGACCGAGCGATCGAGTCATATCCCGGTGTCGCCAGAC 1600
 Db TTCCTACCGCGCGGAGACCGAGCGATCGAGTCATATCCCGGTGTCGCCAGAC 1600
 Oy 352 TATATACCGCGAGATGTAACCAAGCCATAGAGATCGATGATTTCTTTCGCGAAT 411
 Db TATATACCGCGAGATGTAACCAAGCCATAGAGATCGATGATTTCTTTCGCGAAT 411
 Oy 1601 GCCCGCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 1660
 Db GCCCGCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 1660
 Oy 412 CCATCCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 471
 Db CCATCCGCAACATTCTGCTCCCGAGCGAGCTACCAATACAGAGCGCGCGCG 471
 Oy 1661 TTCAACAGGCTGAGAGTCGCGCATTTGCACTCATTCCTCCGACAGAGGGTGGAGATGAC 1720
 Db TTCAACAGGCTGAGAGTCGCGCATTTGCACTCATTCCTCCGACAGAGGGTGGAGATGAC 1720
 Oy 472 TATACCGCGCTCGAATTCGCAAAATACGATCTTCTCCCGAGAGATGGGAATCAAT 531
 Db TATACCGCGCTCGAATTCGCAAAATACGATCTTCTCCCGAGAGATGGGAATCAAT 531


```
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148584P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151030P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157865P.
PR 06-OCT-1999; 99US-0157853P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160811P.
PR 22-OCT-1999; 99US-0160899P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 1.3%; Score 142.2; DB 3; Length 1533;

Best Local Similarity 63.1%; Pred. No. 3.3e-14;

Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```
QY 1481 AGCGCGGTGCGAGAGACCTGTCGCGAGGCGCGCTTACATCTATCGCCGACGACGTC 1540
DB 356 AGGCGCGGTGCGTGAATATTAAACGAGAACTTCCGACGACGCTGAAGCGCCGAGATGTG 415
```

```
QY 1541 TTCTCAACCGCGCGCGGACCCAGGCGATCGAGTCAATAATCCCGGTCTGGCCGAC 1600
DB 416 TATATACCGAGAGATGTATACCAAGCCATAGAGATGTATGATTTCTTGGCGGAAT 475
QY 1601 GCCGCGCAACATTTGTCTCCCGACGAGCTACCAACTACGAGCGCGCGCG 1660
DB 476 CCATCGCGCAACTTTACTTCCAGCGCGGATATCTTCACATGCTGCTGTC 535
QY 1661 TTCAACAGGTGAGGTTCGGCATTTTCACCTTCATCCCGACAGAGGGTGGAGATGAC 1720
DB 536 TTATGCGGCTCGAGATTCGCAAAATACGATCTTCTCCGAGAGTATGGAAATCAAT 595
QY 1721 ATGACCTGCTGGAATTCATTCGCGGCAAGAACACACCGCGCATGGTCATTAACCCC 1780
DB 596 CTGATGCGCTTCGAGCGCGCTGGGTGAGATATACCTGCGATGGTAAATCAACCCC 655
QY 1781 AACCAACCGTGGCGGACGCTTTACTCTTACGACCATCTGTCCAGGT 1827
DB 656 AACCAATCATGTGGAAGAGCTTACCTTACGACCATCTCAAGAGGT 702

RESULT 10
AAC35232
ID AAC35232 standard; DNA; 1535 BP.
XX
AC AAC35232;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9451.
XX
KW Hybridization assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0133407P.
PR 04-MAY-1999; 99US-0133484P.
PR 05-MAY-1999; 99US-0133485P.
PR 06-MAY-1999; 99US-0133486P.
PR 06-MAY-1999; 99US-0133487P.
PR 07-MAY-1999; 99US-0133486P.
PR 11-MAY-1999; 99US-0133425P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
```

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151348P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157717P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158028P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160776P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 1.3%; Score 142.2; DB 3; Length 1535;
 Best Local Similarity 63.1%; Pred. No. 3.3e-14;
 Matches 219; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 1481 AGCGCCGTCGACAGACCTGTGCGAGGCGGTCCGTACCTATTCGCGCCGACGACGTC 1540
 DB 356 AGGGCGGTGCTGATATTAAACGAGAACTTCGACCAAGCGGCGAGATG 415
 QY 1541 TTCCTACCGCGGGGGGACCGAGCGATCGATCATATCCGGTGTGCGCCAGACC 1600
 DB 416 TATATACCGGAGATGTAACCAAGCATAGAGATGATTTCTTCCCGAAT 475
 QY 1601 GCCGCGCCCAATTCGTCTCCCGAGCGGCTACCCAACTACGAGGCGCGCGCG 1660
 DB 476 CCATCGCCCAATTTCTACTTCCAGGCGGGATATCTCTACGATGCTGTCTGTC 535
 QY 1661 TTCAACAGCTGAGAGTCCGCAATTTGACCTTCATCCCGACAGGCGGTGGAGATCGAC 1720
 DB 536 TATAGCGGCTCGAGATTCGCAATACGATCTTCTCCCGAGAGTGTGGAAATCAAT 595
 QY 1721 ATCGACTGGCTGGATTCATTCGCGCAAGAACACACCGCATGTCTATTAACCC 1780
 DB 596 CTCGATGCGCTCGAGGCGGCTCGGATGAGATACCGTCGCAATGATCATCAACCC 655
 QY 1781 AACCAACCGCTGCGGAGCGTTACTCTTACGACCATCTGTCCAGAGT 1827
 DB 656 AACCATTCATGTGAAACGCTTCACTACGACCATCTCAACAGGT 702

RESULT 11

ADA70161
 ID ADA70161 standard; DNA; 930 BP.

XX AC ADA70161;

DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3484.

XX KM Plant; bacterial infection; fungal infection; viral infection; rice;
 gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-1B001105.

XX PR 22-JUN-2001; 2001WO-1B001105.

XX PA (SYN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
 PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175390/17.

XX PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

XX PS Claim 6; SEQ ID NO 3484; 899pp; English.

XX CC The present invention relates to a method (M1) for identifying genes

CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC the incompatible interaction of plant gene expression in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX SQ Sequence 930 BP; 217 A; 259 C; 229 G; 224 T; 0 U; 1 Other;

Query Match 1.3%; Score 141.4; DB 8; Length 930;
 Best Local Similarity 64.1%; Pred. No. 4e-14;
 Matches 230; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 1478 CACAGCGCGTCGACAGACCTGTGCGAGGCGGTCCGTACCTATTCGCGCCGACGAC 1537
 DB 274 CGCGAGCTATTCAGAGTACCTATCTGTATCTTCTTCAACCTTTGACAGATGAT 333
 QY 1538 GTCTTCTTCACCGCGCGGACCGAGCGATCGAGTCAATATCCGGTGTGCGCCAG 1597
 DB 334 ATTTCTTCACATCTGAGAGTACCCACCAATCGAGATGTTATGCTTTTGGCCAA 393
 QY 1598 ACCCGCGCGCAATTCGTCTCCCGAGCGGCTACCCAACTACGAGGCGCGCC 1657
 DB 394 ---CGAGTCCCAATATTTGCTTCCAAAGCCGGGTACCCAAACATGAAAGCATGCG 450
 QY 1658 GCGTTCAACAGCTGAGAGTCCGCAATTCGACCTCATCTCCCGACAGGCGTGGAGATC 1717
 DB 451 GTGTTCCACAGAGTGAAGTGGGCTCATATCTTTGTCAGAGAGAGATGGAGATT 510
 QY 1718 GACATGACTGCTGGAAATTCATTCGCGCAAGAACACACCGCATGTCTATCAATAAC 1777
 DB 511 AATGTGAAGCTGTGAAGCTTTAGCAGATGATATCTGTGCAATGATGATTAATAAC 570
 QY 1778 CCCAACACCGCTGCGGAGGCTTACTCTACGACCATCTGTCCAAAGTTTACATCC 1836
 DB 571 CCCATTAACCTTGTGTATGTGTACCTTATGACATCTGTCCAAATTTGCAATAC 629

RESULT 12

ADJ40562
 ID ADJ40562 standard; CDNA; 930 BP.

XX AC ADJ40562;

DT 06-MAY-2004 (first entry)

XX DE Plant CDNA #1562.

XX KM Plant; gene; ss; transcription; plant genome augmentation; cereal;
 KM soybean; alfalfa; sunflower; cotton; peanut; tobacco; sugar beet;
 KM maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
 KM stress tolerance; salt tolerance; cold tolerance; drought tolerance;
 KM plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
 KM antifungal.

XX OS Eukaryota.

XX PN US2004016025-A1.

XX PD 22-JAN-2004.

XX PF 26-SEP-2002; 2002US-00260238.

XX PR 26-SEP-2001; 2001US-0325277P.

XX PR 26-SEP-2001; 2001US-0325448P.

XX PR 04-APR-2002; 2002US-0370620P.

XX PA (BUDW/) BUDWORTH P.

XX PA (MOUG/) MOUGAMER T.

XX PA (BRIG/) BRIGGS S P.

PA (COOP/) COOPER B.
 PA (GLAZ/) GLAZEBROOK J.
 PA (GOFF/) GOFF S A.
 PA (KATRI/) KATAGIRI F.
 PA (KREP/) KREPS J.
 PA (PROV/) PROVANT N.
 PA (RICK/) RICHE D.
 PA (ZHUT/) ZHU T.
 XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
 PI Goff SA, Katagiri F, Kreps J, Provant N, Riche D, Zhu T;
 XX WPI; 2004-190374/18.
 DR WPI; 2004-190374/18.
 XX New rice promoter, useful for manipulating crop plants to alter or
 PT improve phenotypic characteristics, e.g. produce large quantities of oil
 PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
 PT or high nutritional value.
 PS Claim 25; SEQ ID NO 1562; 230pp; English.
 XX The invention relates to plant nucleotide sequences that direct seed-
 CC leaf-and/or stem-, panicle-, root- or pollen-specific or -preferential
 CC or constitutive transcription of an operatively linked nucleic acid
 CC segment. The invention also relates to a method for augmenting a plant
 CC genome and a method of identifying a gene, where its expression is
 CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
 CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
 CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
 CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
 CC encode are useful for manipulating crop plants to alter or improve
 CC phenotypic characteristics, to produce large quantities of oil or
 CC proteins, to incur resistance to insecticides, viruses or fungi, and to
 CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
 CC have a high nutritional value with reduced apical dominance or dwarfism,
 CC early flowering or altered metabolic pathways. This sequence represents a
 CC plant nucleic acid of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
 XX Sequence 930 BP; 217 A; 259 C; 229 G; 224 T; 0 U; 1 Other;
 SQ
 Query Match 1.3%; Score 141.4; DB 12; Length 930;
 Best Local Similarity 64.1%; Pred. No. 4e-14;
 Matches 230; Conservative 0; Mismatches 126; Indels 3; Gaps 1;
 QY 1478 CACAGCGCGGCGGAGACACCTGTGCGAGGCGGCGGTACATCTATCGGCCGACGAC 1537
 DB 274 CGCCGAGCTATTTGACAGAGTACCTATCTGTGATCTTCTTACAGCTTTGACAGATGAT 333
 QY 1538 GTCTTCTTACCGCGCGGAGACCGAGCGATCGAGTCATATATCCGGTGTGCGCCAG 1597
 DB 334 ATTTTCTTCACTGTGAGGTATCCCAAGCATCGAGATTTGATGTCTGTTTGGCCAA 393
 QY 1598 ACCGCGCGGCGCAACATTTCTGCTCCCAAGCGGCTACCCAACTACGAGCGCGCC 1657
 DB 394 ---CGAGGTGCAATATATTTGCTTCAAAAGCCGGGTACCCAAACATGAGCATCGG 450
 QY 1658 GCGTTCAACAGGCTGGAAGTCCGGCATTTGCACTTCATCCCGGACAAAGGGTGGAGATC 1717
 DB 451 GTGTTCACAGAGATGGAAGTGGGCTCTATATCTTGTTCGAGAAGAGAGATGGAGATT 510
 QY 1718 GACATCGACTGCTGGAATTCATCGCCGACAGAAACACCAACCGCATGGTCTCATTAAC 1777
 DB 511 AATGTGAAGCTGTGGAAGCTTTAGCAATGAGAAATCTGTTCGAAATGATTAATTAAC 570
 QY 1778 CCCAACAACCGGTGCGGAGCGTTTACTCTTACGACCATCTGTCCAAAGTTTCACATCC 1836
 DB 571 CCCAATAACCCCTTGTGTATATGTATGACTTATGAGCATCTGTCCAAAGATTGCAAGATAC 629

ID AAD57651 standard; DNA; 1200 BP.
 XX AAD57651;
 AC AAD57651;
 XX 20-NOV-2003 (first entry)
 DT 20-NOV-2003 (first entry)
 XX Rice enhanced yield gene, NMAAT.
 DE Rice enhanced yield gene, NMAAT.
 XX Rice; abiotic stress tolerance; pathogen resistance; disease resistance;
 KM grain quality; nutritional content; plant yield; NMAAT; plant; gene; ds.
 XX Oryza sativa.
 OS Oryza sativa.
 FH Key Location/Qualifiers
 FH CDS 1..1200
 FT /*tag= a
 FT /product= "Rice NMAAT protein"
 FT
 XX WO2003048319-A2.
 XX 12-JUN-2003.
 PD 12-JUN-2003.
 XX 27-NOV-2002; 2002MO-US038359.
 PF 30-NOV-2001; 2001US-0334501P.
 PR (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA Sainz MB, Salmeron J, Weislo L;
 PI WPI; 2003-505288/47.
 DR P-PSDB; AAE36283.
 XX New nucleic acid from Oryza sativa, useful for altering abiotic stress
 PT tolerance, pathogen or disease resistance or the grain quality,
 PT nutritional content or yield in a plant.
 XX Claim 2; Page 202-203; 223pp; English.
 XX The invention relates to nucleic acid molecules from rice encoding
 CC proteins for abiotic stress tolerance, enhanced pathogen or disease
 CC resistance and altered nutritional quality. The sequences of the
 CC invention are useful for altering abiotic stress tolerance, pathogen or
 CC disease resistance or the grain quality, nutritional content or yield in
 CC a plant. The present sequence is rice enhanced yield gene, NMAAT
 CC
 SQ Sequence 1200 BP; 272 A; 282 C; 354 G; 292 T; 0 U; 0 Other;
 Query Match 1.3%; Score 139.6; DB 10; Length 1200;
 Best Local Similarity 64.1%; Pred. No. 8.6e-14;
 Matches 227; Conservative 0; Mismatches 124; Indels 3; Gaps 1;
 QY 1477 ACAAGAGCGGTGACAGAGACCTGTGCGAGGCGGTGCGATGCTATCGCGCGAGCA 1536
 DB 186 AGGAGGTCTATCCGCGGCTATCTTATCGAGACTTCCATATAGCTATCACTGATGA 245
 QY 1537 GCTTCTTCAACCGCGCGGAGACCGAGCGATCGAGTCATATATCCGGTGTGCGCCA 1596
 DB 246 TGTATCTTGAAGATGAGTGTGCTCAAGGATGAGATCATCTGTCTTACTAGCTCG 305
 QY 1597 GACCGCGGCGCAACATTTCTGCTCCCAAGCGGCTACCCAAACTACAGAGCGCGC 1656
 DB 306 ---CCCTGTGCGCAACATCTCTGTGCCCAAGCGGTACTCTGTTCACAGAGGACGCGC 362
 QY 1657 CGGTTCAACAGGCTGGAAGTCCGGCATTTGCACTTCATCCCGACAAAGGGTGGAGAT 1716
 DB 363 AGTGTCAATGAGCATGAGAGGTCAAGTCTTGTATCTTCTCCAGAGAGTGGCTGGAGAT 422
 QY 1717 CGACATCGACTGCTGGAATTCATCGCCGACAGAAACACCAACCGCATGGTCTCATTAAC 1776
 DB 423 TGATCTTGAAGTGAAGTGAAGAACTTGTGACAAAGACAGGTTGCAATGTCATTAAC 482
 QY 1777 CCCAACAACCGGTGCGGAGCGTTTACTCTTACGACCATCTGTCCAAAGTTTC 1830

Db 483 TCAGAGAAATCCCTGTGGCAATGTACACTTGTGAGCAATTTGGCAAGTTGTC 536

RESULT 14

ADA70274 ID ADA70274 standard; DNA; 1290 BP.

XX ADA70274;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 3597.

XX Plant, bacterial infection; fungal infection; viral infection; rice;

KM gene; ds.

OS Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

PF 22-JUN-2001; 2001WO-1B001105.

PR 22-JUN-2001; 2001WO-1B001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,

PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G,

XX WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to

XX Claim 6; SEQ ID NO 3597; 899bp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 1290 BP; 280 A; 307 C; 400 G; 297 T; 0 U; 6 Other;

XX Query Match 1.3%; Score 137.4; DB 8; Length 1290;

XX Best Local Similarity 60.1%; Pred. No. 2.1e-13; Indels 3; Gaps 1;

XX Matches 247; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

Qy 6962 CGCCGCTGTTCTTCCCGGTGCGTTCAAAATTTTAACTTCTTAATAGTACTTATTAATAA 7021

Db 254 CGGCGCGGTCTCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 313

Qy 7022 CAAACGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 7081

Db 314 CACGAGGCTCTATCGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 373

Qy 7082 ACCTCTTCTCAACCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 7141

Db 374 ATGTGACTGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 433

Qy 7142 AGACTGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 7201

Db 434 GCCCT--GCTGCAACATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 490

Qy 7202 CGGCAATTAACAAGCTGAGGTCCGCACTTCGACTATCCGCCGCAAGGGGTGGAGA 7261

Db 491 CAGTGTTCATATGATGATGAGGTGAGTCAAGTACTTTGATTTTCCAGAGATGCTGGAGG 550

Qy 7262 TCAGATGACATCTGCTGTAATTCATGCGGACAGAACCAACCGGATGATCATATA 7321

Db 551 TTGATCTGATGAGAGTGCAGAACTTGTGACAAAGAACGGTTGGCATGTGCTATATCA 610

Qy 7322 ACCCAACAAATCCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7372

Db 611 ATCCAGAAATCCCTGTGGCAATGTACACTTGTGAGCATTTGGCAAGG 661

RESULT 15

AB214179 ID AB214179 standard; DNA; 1389 BP.

XX AB214179;

XX 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1984.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

XX WO200216655-A2.

PD 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US026685.

PR 24-AUG-2001; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCR1) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and

XX producing plants with increased tolerance to these abiotic stresses.

XX Claim 14; SEQ ID NO 1984; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office

XX Sequence 1389 BP; 389 A; 322 C; 337 G; 341 T; 0 U; 0 Other;

XX Query Match 1.2%; Score 136; DB 6; Length 1389;

XX Best Local Similarity 63.6%; Pred. No. 3.6e-13; Indels 3; Gaps 1;

XX Matches 224; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

Qy 1482 GCGCGGTGCGAGCACTGTGCGAGGGGCTGCGGTAATGTTATGCGCGGCGGAGCTT 1541

Db 344 GAGCGGTGCTGATTATTAAGAACGAGATCTTCCGCAAGATTAACGCTGAAGATATTT 403


```
PF 04-JUL-2000; 2000WO-JP004425.
XX
XX 05-JUL-1999; 99JP-00190318.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Mori S, Nakaniishi H, Takahashi M, Nishizawa N,
XX WPI; 2001-138030/14.
XX
XX Gramineous plant, e.g. rice, with tolerance to iron deficiency for growth
XX in calcareous alkaline soil is constructed by transformation with a gene
XX of encoding an enzyme of the mugineic acids biosynthetic pathway.
XX
XX Example 6; Page 20; 61pp; Japanese.
XX
XX The present invention describes a method for constructing a rice plant
XX with improved iron absorbability and a tolerance to iron deficiency which
XX comprises transferring a gene encoding an enzyme in the mugineic acid
XX biosynthetic pathway into a rice plant. The method is for constructing
XX gramineous plant e.g. rice with tolerance to iron deficiency, which is
XX useful in agriculture in producing new breeds of rice plants capable of
XX vigorous growth in calcareous alkaline soil for improving crop
XX production. The constructed plant has tolerance to iron deficiency, and
XX is therefore capable of vigorous growth in calcareous alkaline soil. The
XX present sequence represents a primer which is used in an example from the
XX present invention
XX
XX Sequence 24 BP; 6 A; 7 C; 5 G; 6 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 24; DB 5; Length 24;
XX Best Local Similarity 100.0%; Pred. No. 0.079; Indels 0; Gaps 0;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TCTGGGATCTCAGTGCATCCACA 24
XX 1 TCTGGGATCTCAGTGCATCCACA 24
XX
XX DB
XX
XX RESULT 2
XX AAF32380
XX ID AAF32380 standard; CDNA; 10966 BP.
XX
XX AAF32380;
XX
XX 18-APR-2001 (first entry)
XX
XX Hordeum vulgare L. var. Igr1 NAAT encoding CDNA SEQ ID NO:1.
XX
XX Hordeum vulgare L. var. Igr1; nicotianamine aminotransferase; NAAT;
XX NAAT-A; NAAT-B; iron deficiency; gramineous plant; barley; rice;
XX mugineic acid biosynthetic pathway; calcareous alkaline soil; ss.
XX
XX Hordeum vulgare.
XX
XX WO200101762-A1.
XX
XX 11-JAN-2001.
XX
XX 04-JUL-2000; 2000WO-JP004425.
XX
XX 05-JUL-1999; 99JP-00190318.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Mori S, Nakaniishi H, Takahashi M, Nishizawa N,
XX WPI; 2001-138030/14.
XX
XX P-PSDB; AAB69048. AAB69049.
XX
XX Gramineous plant, e.g. rice, with tolerance to iron deficiency for growth
XX in calcareous alkaline soil is constructed by transformation with a gene
XX of encoding an enzyme of the mugineic acids biosynthetic pathway.
```

```
XX
XX Claim 6; Fig 10; 61pp; Japanese.
XX
XX The present invention describes a method for constructing a rice plant
XX with improved iron absorbability and a tolerance to iron deficiency which
XX comprises transferring a gene encoding an enzyme in the mugineic acid
XX biosynthetic pathway into a rice plant. The method is for constructing
XX gramineous plant e.g. rice with tolerance to iron deficiency, which is
XX useful in agriculture in producing new breeds of rice plants capable of
XX vigorous growth in calcareous alkaline soil for improving crop
XX production. The constructed plant has tolerance to iron deficiency, and
XX is therefore capable of vigorous growth in calcareous alkaline soil. The
XX present sequence encodes two nicotianamine aminotransferases (NAAT),
XX designated NAAT-A and NAAT-B, isolated from Hordeum vulgare L. var. Igr1
XX (barley), for use in the method of the invention
XX
XX Sequence 10966 BP; 3082 A; 2422 C; 2366 G; 3096 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 24; DB 5; Length 10966;
XX Best Local Similarity 100.0%; Pred. No. 0.24; Indels 0; Gaps 0;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TCTGGGATCTCAGTGCATCCACA 24
XX 8475 TCTGGGATCTCAGTGCATCCACA 8498
XX
XX DB
XX
XX RESULT 3
XX AAS27894/C
XX ID AAS27894 standard; CDNA; 104 BP.
XX
XX AAS27894;
XX
XX 07-NOV-2001 (first entry)
XX
XX Novel CDNA encoding for human respiratory antigen #26.
XX
XX Human; respiratory antigen; respiratory disorder; throat disorder;
XX lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX respiratory active; ss.
XX
XX Homo sapiens.
XX
XX WO200155448-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001333.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
```

Query Match	80.0%	Score 19.2	DB 4	Length 104
Best Local Similarity	87.5%	Pred. No. 23		
Matches	21	Conservative	0	Mismatches 3; Indels 0; Gaps 0;
PR	08-NOV-2000	2000US-0246526P		
PR	08-NOV-2000	2000US-0246527P		
PR	08-NOV-2000	2000US-0246528P		
PR	08-NOV-2000	2000US-0246532P		
PR	08-NOV-2000	2000US-0246539P		
PR	08-NOV-2000	2000US-0246610P		
PR	08-NOV-2000	2000US-0246611P		
PR	08-NOV-2000	2000US-0246613P		
PR	17-NOV-2000	2000US-0249207P		
PR	17-NOV-2000	2000US-0249208P		
PR	17-NOV-2000	2000US-0249209P		
PR	17-NOV-2000	2000US-0249210P		
PR	17-NOV-2000	2000US-0249211P		
PR	17-NOV-2000	2000US-0249212P		
PR	17-NOV-2000	2000US-0249213P		
PR	17-NOV-2000	2000US-0249214P		
PR	17-NOV-2000	2000US-0249215P		
PR	17-NOV-2000	2000US-0249216P		
PR	17-NOV-2000	2000US-0249217P		
PR	17-NOV-2000	2000US-0249218P		
PR	17-NOV-2000	2000US-0249244P		
PR	17-NOV-2000	2000US-0249245P		
PR	17-NOV-2000	2000US-0249264P		
PR	17-NOV-2000	2000US-0249265P		
PR	17-NOV-2000	2000US-0249297P		
PR	17-NOV-2000	2000US-0249299P		
PR	17-NOV-2000	2000US-0249300P		
PR	01-DEC-2000	2000US-0250160P		
PR	01-DEC-2000	2000US-0250391P		
PR	05-DEC-2000	2000US-0251030P		
PR	05-DEC-2000	2000US-0251988P		
PR	05-DEC-2000	2000US-0256719P		
PR	06-DEC-2000	2000US-0251479P		
PR	08-DEC-2000	2000US-0251856P		
PR	08-DEC-2000	2000US-0251868P		
PR	08-DEC-2000	2000US-0251869P		
PR	08-DEC-2000	2000US-0251989P		
PR	08-DEC-2000	2000US-0251990P		
PR	11-DEC-2000	2000US-0254097P		
PR	05-JAN-2001	2001US-0259678P		
PA	(HUMA-) HUMAN GENOME SCI INC.			
PI	Rosen CA, Barash SC, Ruben SM;			
DR	WPI, 2001-476224/51.			
XX	P-PSDB; AAU17710.			
PT	Isolated polypeptide for treating, preventing and/or prognosing			
XX	disorders related to the respiratory system including respiratory cancers			
PT	and also for testing and detection e.g. diagnosis.			
XX				
PS	Claim 4; SED ID No 36; 546P; English.			
XX				
CC	The present invention relates to the isolation of novel human respiratory			
CC	antigens (AAU17685-AAU17975), and cDNA and genomic sequences encoding for			
CC	these polypeptides. The sequences of the invention are useful for			
CC	preventing, treating and/or prognosing disorders related to the			
CC	respiratory system including throat disorders (e.g. vocal cord paralysis,			
CC	tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic			
CC	disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose			
CC	disorders and cancers of the respiratory tissues e.g. lung cancer. The			
CC	polynucleotide sequences of the invention are useful in gene therapy and			
CC	antisense therapy. AAS27869-AAS28159 encode for novel human respiratory			
CC	antigens. Note: The sequence data for this patent did not form part of			
CC	the printed specification, but was obtained in electronic format directly			
XX	from WIPO at ftp.wipo.int/pub/published_pct_sequences			
XX				
SO	Sequence 104 BP; 32 A; 19 C; 23 G; 29 T; 0 U; 1 Other;			

OY 1 TCTGGATCTCAGTCATCCACA 24
 DB 68 TCTGGCTTTCAGTCATCCATCA 45

RESULT 4
 ADG40798/c
 ID ADG40798 standard; cDNA; 104 BP.
 XX ADG40798;
 AC
 DT 26-FEB-2004 (fixat entry)
 XX
 DE Human respiratory system associated protein cDNA seq id 36.
 XX
 KW antiinflammatory; antiallergic; antiaesthetic; cytostatic; gene therapy;
 KW respiratory system antigen;
 KW human respiratory system associated polynucleotide;
 KW respiratory system disorder; throat disorder; vocal cord paralysis;
 KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
 KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
 KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
 KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
 KW cancer of the nose; gene therapy; chromosome identification; forensic;
 KW human respiratory system associated protein; gene; ss; human.
 XX
 OS Homo sapiens.
 XX
 PN US2003215893-A1.
 XX
 PD 20-NOV-2003.
 XX
 PF 07-AUG-2002; 2002US-00212872.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225477P.
 PR 14-AUG-2000; 2000US-0225575P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-022679P.
 PR 22-AUG-2000; 2000US-022681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0235484P.
 PR 26-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 28-SEP-2000; 2000US-0235935P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-024474P.
 PR 08-NOV-2000; 2000US-024475P.
 PR 08-NOV-2000; 2000US-024476P.
 PR 08-NOV-2000; 2000US-024477P.
 PR 08-NOV-2000; 2000US-024527P.
 PR 08-NOV-2000; 2000US-024528P.
 PR 08-NOV-2000; 2000US-024529P.
 PR 08-NOV-2000; 2000US-024532P.
 PR 08-NOV-2000; 2000US-0245324P.
 PR 08-NOV-2000; 2000US-0245524P.
 PR 08-NOV-2000; 2000US-0245525P.
 PR 08-NOV-2000; 2000US-0245526P.
 PR 08-NOV-2000; 2000US-024611P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.

PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236388P.
PR 29-SEP-2000; 2000US-0236399P.
PR 29-SEP-2000; 2000US-0236399P.
PR 29-SEP-2000; 2000US-0236399P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 11-DEC-2000; 2000US-0251909P.
PR 05-JAN-2001; 2000US-0254097P.
PR 17-JAN-2001; 2000US-0259678P.
PR 17-JAN-2001; 2000US-00764860.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Ruben SM, Barash SC;
XX
XX DR WPI, 2003-765403/72.
XX DR E-PSDB; ADI96864.
XX
XX PT New human respiratory system-related polypeptide and genes, useful for
XX PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
XX PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
XX PT sinusitis.
XX
XX PS Claim 1; SEQ ID NO 36; 202pp; English.
XX
XX CC This invention is related to a novel isolated polypeptide, which
XX CC comprises a human respiratory system-related polypeptide, and the DNA
XX CC sequence which encodes it. The invention may be useful for the
XX CC development of compounds with an antiasthmatic, antibacterial,
XX CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
XX CC addition, the sequences disclosed may be useful for gene therapy. The
XX CC polypeptide or polynucleotide is useful for treating, preventing or
XX CC ameliorating a medical condition, for example pneumonia, lung cancer,
XX CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukemia,
XX CC inflammations, sinusitis, chronic obstructive pulmonary disease or
XX CC infectious diseases. The polypeptide or polynucleotide is also useful for
XX CC diagnosing any of these diseases or a susceptibility to the disease. The
XX CC present sequence is that of a respiratory system associated human gene of
XX CC the invention.
XX
XX SQ Sequence 104 BP; 32 A; 19 C; 23 G; 29 T; 0 U; 1 Other;
XX
XX Query Match 80.0%; Score 19.2; DB 11; Length 104;
XX Best Local Similarity 87.5%; Pred. No. 23;
XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 TCTGGATCTCAGTCATCCACA 24
XX |||||
XX DB 68 TCTGGGCTTCAGTCATCCATCA 45
XX
XX RESULT 6
XX AAS28222/c
XX ID AAS28222 standard; DNA; 224 BP.
XX
XX AC AAS28222;
XX
XX DT 07-NOV-2001 (first entry)
XX
XX DE Genomic sequence #62 encoding for novel human respiratory antigen.
XX
XX KW Human; respiratory antigen; respiratory disorder; throat disorder;
XX KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX KW respiratory active; ds.
XX
XX OS Homo sapiens.
XX
XX XX

PN WO200155448-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001333.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186550P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205151P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225577P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231444P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254907P.
PR 05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX
PI

XX WPI, 2001-476224/51.
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the respiratory system including respiratory cancers
XX and also for testing and detection e.g. diagnosis.
PS Disclosure, SEQ ID No 656, 546bp, English.
XX
XX The present invention relates to the isolation of novel human respiratory
XX antigens (AAU17685-AAU17975), and cDNA and genomic sequences encoding for
XX these polypeptides. The sequences of the invention are useful for
XX preventing, treating and/or prognosing disorders related to the
XX respiratory system including throat disorders (e.g. vocal cord paralysis,
XX tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
XX disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
XX disorders and cancers of the respiratory tissues e.g. lung cancer. The
XX polynucleotide sequences of the invention are useful in gene therapy and
XX antineoplastic therapy. AAS28161-AAS28764 represent genomic sequences encoding
XX for novel human respiratory antigens. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 224 BP; 76 A; 39 C; 58 G; 51 T; 0 U; 0 Other;
Query Match 80.0%; Score 19.2; DB 4; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TCTGGATCTCAGTCATCCACA 24
Db 162 TCTGGCTTTCAGTCATCCATCA 139
RESULT 7
AAS28221/c
ID AAS28221 standard; DNA; 224 BP.
XX
XX AAS28221;
XX
XX 07-NOV-2001 (first entry)
XX
XX Genomic sequence #61 encoding for novel human respiratory antigen.
DE
XX Human; respiratory antigen; respiratory disorder; throat disorder;
XX lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX respiratory active; ds.
XX
XX Homo sapiens.
XX
XX MO200155448-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US001333.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229345P.
PR 01-SEP-2000; 2000US-0229349P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.

PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256713P.
PR 06-DEC-2000; 2000US-0251472P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0253678P.
PR 17-JAN-2001; 2001US-00764860.
PR 14-FEB-2002; 2002US-00074095.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-902033/82.
XX
XX Novel respiratory system antigen and polynucleotides encoding the
PT polypeptides, useful for treating diagnosing, treating or preventing
PT consiliitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
PT cancer.
XX
PS Disclosure; SEQ ID NO 657; 236pp; English.
XX
XX The invention describes an isolated polypeptide (I) comprising an amino
CC acid sequence that is at least 90% identical to polypeptide fragment of
CC any one of 299 respiratory system antigen sequences (PS) and having
CC biological activity, polypeptide domain or epitope of PS, full-length
CC protein of PS, or variant allelic variant or species homolog of PS. (I)
CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition in
CC a subject which involves determining the presence or absence of mutation
CC in (II) or determining the presence or amount of expression of (I) in a
CC biological sample and diagnosing a pathological condition based on the
CC result. The human respiratory system associated polynucleotides, the
CC polypeptides encoded by them, and antibodies that immunospecifically bind
CC these polypeptides are useful in diagnosis, treatment, prevention and/or
CC prognosis of disorders of respiratory system such as throat disorders
CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
CC barcodosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
CC cancer of the nose). The polynucleotides are useful in gene therapy
CC techniques, for chromosome identification, identifying individuals from

Query Match 80.0%; Score 19.2; DB 10; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCTGGATCTCAGTCATCAACA 24
162 TCTGGCTTTCATGCATCATCA 139
Db

RESULT 10
ADG41418/c
ID ADG41418 standard; DNA; 224 BP.
XX
AC ADG41418;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human respiratory system associated genomic DNA seq id 656.
XX
KW antiinflammatory; antiallergic; antiasthmatic; cyrostatic; gene therapy;
KW respiratory system antigen;
KW human respiratory system associated polynucleotide;
KW respiratory system disorder; throat disorder; vocal cord paralysis;

KW consiliitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KW histiocytosis; barcodosis; nose disorder; rhinitis; sinusitis; neoplasm;
KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
KW cancer of the nose; gene therapy; chromosome identification; forensic;
XX human respiratory system associated protein; ds; human.
OS Homo sapiens.
XX
XX US2003215893-A1.
PD
PD 20-NOV-2003.
XX
XX 07-AUG-2002; 2002US-00212872.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.

```

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0235935P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0241222P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241825P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246522P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.

```

```

PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-0076486P.
PR 14-FEB-2002; 2002US-00074095.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WPI, 2003-902033/82.
XX
XX Novel respiratory system antigen and polynucleotides encoding the
XX polypeptides, useful for treating diagnosing, treating or preventing
XX tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
XX cancer.
XX
XX Disclosure; SEQ ID NO 656; 236bp; English.
XX
XX The invention describes an isolated polypeptide (I) comprising an amino
XX acid sequence that is at least 90% identical to polypeptide fragment of
XX any one of 299, respiratory system antigen sequences (PS) and having
XX biological activity, polypeptide domain or epitope of PS, full-length
XX protein of PS, or variant, allelic variant or species homolog of PS. (I)
XX or a polynucleotide (II) encoding (I) is also useful for diagnosing a
XX pathological condition or a susceptibility to a pathological condition in
XX a subject which involves determining the presence or absence of mutation
XX in (II) or determining the presence or amount of expression of (I) in a
XX biological sample and diagnosing a pathological condition based on the
XX result. The human respiratory system associated polynucleotides, the
XX polypeptides encoded by them, and antibodies that immunospecifically bind
XX these polypeptides are useful in diagnosis, treatment, prevention and/or
XX prognosis of disorders of respiratory system such as throat disorders
XX (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
XX (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
XX pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
XX sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
XX cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
XX cancer of the nose). The polynucleotides are useful in gene therapy
XX techniques, for chromosome identification, identifying individuals from
XX
XX
XX Query Match 80.0%; Score 19.2; DB 10; Length 224;
XX Best Local Similarity 87.5%; Pred. No. 27;
XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 TCTGGATCTCAGTCATCCACA 24
XX |||||
XX Db 162 TCTGGGCTTCAGTCATCCATCA 139
XX
XX RESULT 11
XX ADG41417/c
XX ID ADG41417 standard; DNA; 224 BP.
XX
XX AC ADG41417;
XX
XX DT 26-FEB-2004 (first entry)
XX
XX DE Human respiratory system associated genomic DNA seq id 655.
XX
XX XX antiinflammatory; antiallergic; antiaesthetic; cytostatic; gene therapy;
XX human respiratory system antigen;
XX human respiratory system associated polynucleotide;
XX respiratory system disorder; throat disorder; vocal cord paralysis;
XX tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
XX asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
XX histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
XX cancer; respiratory tissue cancer; throat cancer; lung cancer;
XX cancer of the nose; gene therapy; chromosome identification; forensic;
XX human respiratory system associated protein; ds; human.
XX
XX

```

OS Homo sapiens.
 XX
 PN US2003215893-A1.
 XX
 PD 20-NOV-2003.
 XX
 PF 07-AUG-2002; 2002US-00212872.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198122P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0228287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 12-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 28-SEP-2000; 2000US-0235935P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246533P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249287P.
 PR 17-NOV-2000; 2000US-0249289P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764860.
 PR 14-FEB-2002; 2002US-00074095.

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-902033/82.
 XX Novel respiratory system antigen and polynucleotides encoding the
 PT polypeptides, useful for treating diagnosing, treating or preventing
 PT tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
 cancer.
 XX Disclosure; SEQ ID NO 655; 236pp; English.
 XX The invention describes an isolated polypeptide (I) comprising an amino
 CC acid sequence that is at least 90% identical to polypeptide fragment of
 CC any one of 239 respiratory system antigen sequences (PS) and having
 CC biological activity, polypeptide domain or epitope of PS, full-length
 CC protein of PS, or variant, allelic variant or species homolog of PS. (I)
 CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
 CC pathological condition or a susceptibility to a pathological condition in
 CC a subject which involves determining the presence or absence of mutation
 CC in (II) or determining the presence or amount of expression of (I) in a
 CC biological sample and diagnosing a pathological condition based on the
 CC result. The human respiratory system associated polynucleotides, the
 CC polypeptides encoded by them, and antibodies that immunospecifically bind
 CC these polypeptides are useful in diagnosis, treatment, prevention and/or
 CC prognosis of disorders of respiratory system such as throat disorders
 CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
 CC (e.g., pneumonia), allergic disorders (e.g., asthma and eosinophilic
 CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
 CC sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
 CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
 CC cancer of the nose). The polynucleotides are useful in gene therapy
 CC techniques, for chromosome identification, identifying individuals from

Query Match 80.0%; Score 19.2; DB 10; Length 224;
 Best Local Similarity 87.5%; Pred. No. 27;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCTGGATCTCAGTCATCCACA 24
 162 TCTGGCTTCTCAGTCATCCACA 139

RESULT 12
 AD197191/c
 ID AD197191 standard; DNA; 224 BP.
 XX
 AC AD197191;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Human respiratory system associated polypeptide-related DNA SegID55.
 XX
 KW respiratory system-related polypeptide; antiasthmatic; antibacterial;
 KW antiinflammatory; cytotoxic; antianaemic; antiallergic; gene therapy;
 KW pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;
 KW anaemia; leukaemia; inflammation; sinusitis;
 KW chronic obstructive pulmonary disease; infectious disease; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN US200307704-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 14-FEB-2002; 2002US-00074095.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217486P.
 PR 14-JUL-2000; 2000US-0218293P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 14-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 25-SEP-2000; 2000US-0234984P.
 PR 26-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 27-SEP-2000; 2000US-0235837P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236328P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Ruben SM, Barash SC;
XX MPI; 2003-765403/72.
XX
XX New human respiratory system-related polypeptide and genes, useful for
PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
PT sinusitis.
PS Disclosure; SEQ ID NO 655; 202bp; English.

XX This invention is related to a novel isolated polypeptide, which
CC comprises a human respiratory system-related polypeptide, and the DNA
CC sequence which encodes it. The invention may be useful for the
CC development of compounds with an antiasthmatic, antibacterial,
CC antiinflammatory, cytosstatic, antianemic or antiallergic activity. In
CC addition, the sequences disclosed may be useful for gene therapy. The
CC polypeptide or polynucleotide is useful for treating, preventing or
CC ameliorating a medical condition, for example pneumonia, lung cancer,
CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anemia, leukemia,
CC inflammations, sinusitis, chronic obstructive pulmonary disease or
CC infectious diseases. The polypeptide or polynucleotide is also useful for
CC diagnosing any of these diseases or a susceptibility to the disease. The
CC present sequence is that of a human DNA sequence which is related to a
XX human respiratory system associated gene of the invention.
SQ Sequence 224 BP; 76 A; 39 C; 58 G; 51 T; 0 U; 0 Other;
Query Match 80.0%; Score 19.2; DB 11; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCTGGGATCTCAGTCATCCACA 24
Db 162 TCTGGGCTTTCAGTCATCCATCA 139
RESULT 13
AD197192/C
ID AD197192 standard; DNA; 224 BP.
XX
AC AD197192;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human respiratory system associated polypeptide-related DNA SeqID656.
XX
KW antinflammatory; cytostratic; antianemic; antiallergic; gene therapy;
KW pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;
KW anemia; leukemia; inflammatory disease; sinusitis;
XX chronic obstructive pulmonary disease; infectious disease; human; ds.
OS Homo sapiens.
XX
PN US200307704-A1.
XX
PD 24-APR-2003.
XX
PF 14-FEB-2002; 2002US-00074095.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 28-JUN-2000; 2000US-0209467P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.

PR	08-NOV-2000;	2000US-02465525P.
PR	08-NOV-2000;	2000US-02465526P.
PR	08-NOV-2000;	2000US-02465527P.
PR	08-NOV-2000;	2000US-02465528P.
PR	08-NOV-2000;	2000US-02465532P.
PR	08-NOV-2000;	2000US-02465609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250301P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254077P.
PR	05-JAN-2001;	2001US-0259678P.
PR	17-JAN-2001;	2001US-00764860.
PA	(HOMA-) HUMAN GENOME SCT INC.	
PL	Rosen CA, Ruben SM, Barash SC;	
XX		
DR	WPI; 2003-765403/72.	
PT	New human respiratory system-related polypeptide and genes, useful for	
PT	treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic	
PT	fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or	
PT	sinusitis.	
XX		
P5	Disclosure; SEQ ID NO 656; 202pp; English.	
XX		
CC	This invention is related to a novel isolated polypeptide, which	
CC	comprises a human respiratory system-related polypeptide, and the DNA	
CC	sequence which encodes it. The invention may be useful for the	
CC	development of compounds with an antiasthmatic, antibacterial,	
CC	antiinflammatory, cytostatic, antianemic or antiallergic activity. In	
CC	addition, the sequences disclosed may be useful for gene therapy. The	
CC	polypeptide or polynucleotide is useful for treating, preventing or	
CC	ameliorating a medical condition, for example pneumonia, lung cancer,	
CC	cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,	
CC	inflammations, sinusitis, chronic obstructive pulmonary disease or	
CC	infectious diseases. The polypeptide or polynucleotide is also useful for	
CC	diagnosing any of these diseases or a susceptibility to the disease. The	
CC	present sequence is that of a human DNA sequence which is related to a	
CC	human respiratory system associated gene of the invention.	
XX		
SQ	Sequence 224 BP; 76 A; 39 C; 58 G;51 T; 0 U; 0 Other;	
XX		

PR 01-SEP-2000; 2000US-0229345P

[illegible]

PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249257P.
 PR 17-NOV-2000; 2000US-0249259P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250150P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764860.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-765403/72.
 DR
 XX
 PT New human respiratory system-related polypeptide and genes, useful for
 PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
 PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
 PT sinusitis.
 XX
 XX
 PS Disclosure; SEQ ID NO 657; 202pp; English.
 XX
 XX
 CC This invention is related to a novel isolated polypeptide, which
 CC comprises a human respiratory system-related polypeptide, and the DNA
 CC sequence which encodes it. The invention may be useful for the
 CC development of compounds with an antiasthmatic, antibacterial,
 CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
 CC addition, the sequence disclosed may be useful for gene therapy. The
 CC polypeptide or polynucleotide is useful for treating, preventing or
 CC ameliorating a medical condition, for example pneumonia, lung cancer,
 CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
 CC inflammations, sinusitis, chronic obstructive pulmonary disease or
 CC infectious diseases. The polypeptide or polynucleotide is also useful for
 CC diagnosing any of these diseases or a susceptibility to the disease. The
 CC present sequence is that of a human DNA sequence which is related to a
 CC human respiratory system associated gene of the invention.
 CC
 CC
 XX
 SQ Sequence 224 BP; 76 A; 39 C; 58 G; 51 T; 0 U; 0 Other;
 Query Match 80.0%; Score 19.2; DB 11; Length 224;
 Best Local Similarity 87.5%; Pred. No. 27;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 TCTGGGATCTCAGTCATCCACA 24
 DB 162 TCTGGGCTTCTCAGTCATCCATCA 139

RESULT 15
 AAQ48675/c
 ID AAQ48675 standard; cDNA; 3323 BP.
 XX
 XX AAQ48675;
 AC
 XX
 XX
 DT 25-MAR-2003 (revised)
 DT 25-FEB-1994 (first entry)
 XX

DE Human myotonic dystrophy (DM) gene.
 XX
 XX Abnormality; muscular dystrophy; CHR 19; chromosome 19; protein kinase;
 KM adult; ss.
 XX
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..1749
 FT /tag= a
 FT /note= "coding region"
 FT 1972..2004
 FT /tag= b
 FT /note= "CTG repeat region"
 XX
 PN W09317104-A1.
 XX
 PD 02-SEP-1993.
 XX
 XX 19-FEB-1993; 93WO-US001545.
 PF
 XX 20-FEB-1992; 92US-00839255.
 PR
 XX (MAST) MASSACHUSETTS INST TECHNOLOGY.
 PA
 XX Brook JD, Houseman DE;
 PI
 XX
 DR WPI; 1993-288410/36.
 DR P-FSDB; AAR41000.
 XX
 XX
 PT DNA sequence of myotonic dystrophy gene - used to produce probes and
 PT identify CHR 19 abnormality and protein kinase responsible.
 PS
 XX Claim 3; Fig 5; 64pp; English.
 XX
 CC The sequence is that of the adult human brain cDNA clone C28 which
 CC comprises the "myotonic dystrophy gene" which includes in its 3'
 CC untranslated region a transcribed CTG triplet repeat. The sequence
 CC encodes a protein kinase, detection of expression of this protein kinase
 CC may be used as a method for determining whether an individual is likely
 CC to be affected with myotonic dystrophy. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 CC
 CC
 XX
 SQ Sequence 3323 BP; 590 A; 1021 C; 1062 G; 608 T; 0 U; 42 Other;
 Query Match 80.0%; Score 19.2; DB 2; Length 3323;
 Best Local Similarity 87.5%; Pred. No. 44;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 TCTGGGATCTCAGTCATCCACA 24
 DB 3222 TCGGGGGTCTCAGTCATCCAAA 3199

Search completed: October 15, 2005, 02:56:33
 Job time : 14.8797 secs

	10966 bp	DNA	linear	PAT 31-JAN-2002
Creation of iron-deficiency resistant rice plant.				
BA8893				
BA8893.1	GI:18629372			
JP 2001017012-A/1.				
unidentified				
unclassified.				
1 (bases 1 to 10966)				
Mori, S., Nakanishi, N., Takahashi, M. and Nishizawa, N.				
Creation of iron-deficiency resistant rice plant				
Patent: JP 2001017012-A 1 23-JAN-2001;				
SCIENCE & TECH AGENCY, SATOSHI MORI				
OS Hordeum vulgare L. var. Igri				
PN JP 2001017012-A/1				
PD 23-JAN-2001				
PF 05-JUL-1999 JP 1999190318				
PR SATOSHI MORI, NOBUHITO NAKANISHI, MICHIO TAKAHASHI, PI				
NISHIZAWA				
PC A01H/00, C12N5/10, C12N15/09//C12N9/10, C12N15/00 CC				
FH Key	Location/Qualifiers			
FT source	1.10966			
	Location/Qualifiers			
	1.10966			
	Location/Qualifiers			
	1.10966			
	/organism="unidentified"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:32644"			
100.0%; Score 10966; DB 6; length 10966;				
100.0%; Pred. No. 0;				
10966; Conservative 0; Mismatches 0; Indels 0; Gaps 0				
1 CTCGATCCCATGGCAATGTTATGATGCTATCAAGCAAGAAAGAGATGGCATGTGCC 60				
1 CTCGATCCCATGGCAATGTTATGATGCTATCAAGCAAGAAAGAGATGGCATGTGCC 60				
1 CTGTGTGTGATCCCTCACTGCGCTTGGCGCAATGGCGATACCGAGTTAGGTAGAGTGTTT 120				

Db 61 CTGTGTGTCATCCCTCACTGGCTTGGCGAATGGCATCCGAGTTAGTAGAGTTT 120
 Qy 121 TTACATGATGTCGCGCGGCACTGCGAAGAAACTCGTGAGGGAAGTGAAGAGTT 180
 Db 121 TTACATGATGTCGCGCGGCACTGCGAAGAAACTCGTGAGGGAAGTGAAGAGTT 180
 Qy 181 GAGCGATGATGTCGTCGTGATGAGCGAGCTGAGTGGTGTCACTAACTGAACCAATCA 240
 Db 181 GAGCGATGATGTCGTCGTGATGAGCGAGCTGAGTGGTGTCACTAACTGAACCAATCA 240
 Qy 241 GCATTTGGTGTGAGAGTCGAGAGACATGCTTTCGTGCTGCCGATCCGCTTATCTTT 300
 Db 241 GCATTTGGTGTGAGAGTCGAGAGACATGCTTTCGTGCTGCCGATCCGCTTATCTTT 300
 Qy 301 TTCTCCCAATTTATTAAGAGGATAGATGATGGTGTCTGGTTGGGTAGAGTACCTGC 360
 Db 301 TTCTCCCAATTTATTAAGAGGATAGATGATGGTGTCTGGTTGGGTAGAGTACCTGC 360
 Qy 361 ATAGAACCAAGCGAGCGCGCAAAATATGCGGGGATATATGTGGCAGCGCCGCAACGCG 420
 Db 361 ATAGAACCAAGCGAGCGCGCAAAATATGCGGGGATATATGTGGCAGCGCCGCAACGCG 420
 Qy 421 CAGCCCGCTGACGTCGCGAGCGCGCTGCCAGAGCGTGGCCAGCGTGGCGTGGTCT 480
 Db 421 CAGCCCGCTGACGTCGCGAGCGCGCTGCCAGAGCGTGGCCAGCGTGGCGTGGTCT 480
 Qy 481 TCTTGTGTCGCGCGCGCTGCTGCTGTCGCTGAGAGCAACGCGCTATATAGACCGTCAATC 540
 Db 481 TCTTGTGTCGCGCGCGCTGCTGCTGTCGCTGAGAGCAACGCGCTATATAGACCGTCAATC 540
 Qy 541 ACCGCTACTCAATCCGTCCCACTCGTTCTATATACCGCTAATAAGATATTCGTGT 600
 Db 541 ACCGCTACTCAATCCGTCCCACTCGTTCTATATACCGCTAATAAGATATTCGTGT 600
 Qy 601 GTATGTAGATGATCT 660
 Db 601 GTATGTAGATGATCT 660
 Qy 661 ACAGCGAAGCGAGCGAGTCCGCGGAAACGCGCTTGGCGGCGAGCGCGGAAACGCGCAA 720
 Db 661 ACAGCGAAGCGAGCGAGTCCGCGGAAACGCGCTTGGCGGCGAGCGCGGAAACGCGCAA 720
 Qy 721 GAGCAACGCGCATGCGTGTGCTGCGCTGTGAACGCGAAGCAACGCGCATGCGTGTGA 780
 Db 721 GAGCAACGCGCATGCGTGTGCTGCGCTGTGAACGCGAAGCAACGCGCATGCGTGTGA 780
 Qy 781 TGCCGACGCGCAACGCGCAAGAGCAACGCGCATGCGTGTGCTGCGCAACGCGCAAGAG 840
 Db 781 TGCCGACGCGCAACGCGCAAGAGCAACGCGCATGCGTGTGCTGCGCAACGCGCAAGAG 840
 Qy 841 CAACGCGCATGCGAGGCACTGCGAAGCGGCAACGCGGAGGCACTGCGAAGCGGCAAGAG 900
 Db 841 CAACGCGCATGCGAGGCACTGCGAAGCGGCAACGCGGAGGCACTGCGAAGCGGCAAGAG 900
 Qy 901 CAACGCGCATGCGAGGCAACGCGCATGCTGAGGCGCGCAACGCGAAGCGGCAAGAGCA 960
 Db 901 CAACGCGCATGCGAGGCAACGCGCATGCTGAGGCGCGCAACGCGAAGCGGCAAGAGCA 960
 Qy 961 CGAGCATGCGAGGCACTCCGCGGCGAAGCGGCAAGAGCAACGCGGATCCGCGGCGGCG 1020
 Db 961 CGAGCATGCGAGGCACTCCGCGGCGAAGCGGCAAGAGCAACGCGGATCCGCGGCGGCG 1020
 Qy 1021 AGAGGAGAGAGGCGGTCGAGTGAATTTTCGCGGTCGCAAGAGCGGCGTGTGCGGCG 1080
 Db 1021 AGAGGAGAGAGGCGGTCGAGTGAATTTTCGCGGTCGCAAGAGCGGCGTGTGCGGCG 1080
 Qy 1081 GACGCGGCGCAACATGAGCATCCGCGGCGATTCGCGTCAAGATCAGCGGCGGCGAGGA 1140
 Db 1081 GACGCGGCGCAACATGAGCATCCGCGGCGATTCGCGTCAAGATCAGCGGCGGCGAGGA 1140
 Qy 1141 GAAGGCGCGCGCGCGTGTGCTGCGCTGCGCCAGGCGGACCCGTCGCTTCCGCGCTT 1200
 Db 1141 GAAGGCGCGCGCGCGTGTGCTGCGCTGCGCCAGGCGGACCCGTCGCTTCCGCGCTT 1200

Db 1141 GAAGGCGCGCGCGCGTGTGCTGCGCTGCGCCAGGCGGACCCGTCGCTTCCGCGCTT 1200
 Qy 1201 CCGGACGCGCGTGTGAGGCGCGAGAGCGCGCTGCGCGCGCGCTGCGACCGCGCACTTCAA 1260
 Db 1201 CCGGACGCGCGTGTGAGGCGCGAGAGCGCGCTGCGCGCGCGCTGCGACCGCGCACTTCAA 1260
 Qy 1261 CTGCTACCCCGCGCGCTGCGCGCTCCCGCGGCAAGAGGTAAACAAACAAACAAACAA 1320
 Db 1261 CTGCTACCCCGCGCGCTGCGCGCTCCCGCGGCAAGAGGTAAACAAACAAACAAACAA 1320
 Qy 1321 GAACAAATTTCTTTTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
 Db 1321 GAACAAATTTCTTTTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
 Qy 1381 TTCAAGTGTCCGTCCGTCGCTGCAACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
 Db 1381 TTCAAGTGTCCGTCCGTCGCTGCAACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
 Qy 1441 GACCTTCTCCACCTTATATACCAACAAACAAACAAACAAACAAACAAACAAACAAACCT 1500
 Db 1441 GACCTTCTCCACCTTATATACCAACAAACAAACAAACAAACAAACAAACAAACAAACCT 1500
 Qy 1501 GTGCGAGGCGGTGCGGTCATGCTATGCGGCGAGAGCTTCTCTCTCTCTCTCTCTCTCT 1560
 Db 1501 GTGCGAGGCGGTGCGGTCATGCTATGCGGCGAGAGCTTCTCTCTCTCTCTCTCTCTCT 1560
 Qy 1561 CCAAGCGATGAGAGTCAATATCCCGGTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1620
 Db 1561 CCAAGCGATGAGAGTCAATATCCCGGTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1620
 Qy 1621 CCCGAGGCGAGGTCACCAACCTACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
 Db 1621 CCCGAGGCGAGGTCACCAACCTACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
 Qy 1681 GCAATTTGACCTGATCCCGGCAAGAGGCGGAGATGAGATGAGATGAGATGAGATGAGAT 1740
 Db 1681 GCAATTTGACCTGATCCCGGCAAGAGGCGGAGATGAGATGAGATGAGATGAGATGAGAT 1740
 Qy 1741 CGCGCAACAGAACACACCGCATGCTCATATAACCCCAACAAACCGTGGCGGAGCGT 1800
 Db 1741 CGCGCAACAGAACACACCGCATGCTCATATAACCCCAACAAACCGTGGCGGAGCGT 1800
 Qy 1801 TTAATCTTACGACATCTGTCCAAAGTTTCAATCTTGTGCTGTAATGATGATCA 1860
 Db 1801 TTAATCTTACGACATCTGTCCAAAGTTTCAATCTTGTGCTGTAATGATGATGATCA 1860
 Qy 1861 GTTCAATGACCTGTGCAATCTTTTGTCCAAATGCAATGCAATGCAATGCAATGCAATTA 1920
 Db 1861 GTTCAATGACCTGTGCAATCTTTTGTCCAAATGCAATGCAATGCAATGCAATGCAATTA 1920
 Qy 1921 GGTGCGGAGGTGCGGAAAGGCTCGGAAATATGCTGATGCTGACGAGGATATACGCA 1980
 Db 1921 GGTGCGGAGGTGCGGAAAGGCTCGGAAATATGCTGATGCTGACGAGGATATACGCA 1980
 Qy 1981 GCTGTGTCGCGAGCGCGCGCTTCAATCCCAAGGAGTGTGGGCAACATCACCCCTGT 2040
 Db 1981 GCTGTGTCGCGAGCGCGCGCTTCAATCCCAAGGAGTGTGGGCAACATCACCCCTGT 2040
 Qy 2041 GCTGTGCAATAGGTCCTGTGCAAGTCATGAGATGAGTGGCTGAGTGGCTGAGTGGGT 2100
 Db 2041 GCTGTGCAATAGGTCCTGTGCAAGTCATGAGATGAGTGGCTGAGTGGCTGAGTGGGT 2100
 Qy 2101 AGCGGTGTACGACCCGAGAAAGATCTTACAGAAACCTAAGTACTTAAATCTCTATATCA 2160
 Db 2101 AGCGGTGTACGACCCGAGAAAGATCTTACAGAAACCTAAGTACTTAAATCTCTATATCA 2160
 Qy 2161 TTCTTTTCAATGCTAATAGTATTAATAGTACTGTAACATATATTTGCTAAT 2220
 Db 2161 TTCTTTTCAATGCTAATAGTATTAATAGTACTGTAACATATATTTGCTAAT 2220
 Qy 2221 TTGTACTGACATTTTGT 2280
 Db 2221 TTGTACTGACATTTTGT 2280

Oy	2281	ANAGCCGACCACTTCACTTCAAGTCAGTCTTTGGTATTTACTCTGTTTCAAGAAATTA	2340
Dp	2281	CAGACCCAGCAACCTTCACTTCAAGTCAGTCTTTGGTATTTACTCTGTTTCAAGAAATTA	2340
Oy	2341	GTCCTTGGTATTTACTCCTCCTGTCCTAATTTTGGTCGCGTCCCTATGTTAGAGCC	2400
Dp	2341	GTCCTTGGTATTTACTCCTCCTGTCCTAATTTTGGTCGCGTCCCTATGTTAGAGCC	2400
Oy	2401	CACGTGCATGTCMAAGTACCGTTTTTTCACTTAAGTTGAAAGTCMAAGTCACACAT	2460
Dp	2401	CACGTGCATGTCMAAGTACCGTTTTTTCACTTAAGTTGAAAGTCMAAGTCACACAT	2460
Oy	2461	ACACTTGGAGTTAATTTTACTCTTGTGTTGCTTGTGATCCGATAAATTAATAATCAAAA	2520
Dp	2461	ACACTTGGAGTTAATTTTACTCTTGTGTTGCTTGTGATCCGATAAATTAATAATCAAAA	2520
Oy	2521	CTGAACCTACTGTTGAATATAACACTGTTCTTACAGATATACATGATGCACTATGG	2580
Dp	2521	CTGAACCTACTGTTGAATATAACACTGTTCTTACAGATATACATGATGCACTATGG	2580
Oy	2581	CATGCAATATCTTTTGGGTCAAGTACAGATAGTGGAACTCTTTTGAATAATGAT	2640
Dp	2581	CATGCAATATCTTTTGGGTCAAGTACAGATAGTGGAACTCTTTTGAATAATGAT	2640
Oy	2641	ACATTGTACTATGAGTATACCAATTTTATTAAGAAATTTACATATTTTGATATCTTGATGT	2700
Dp	2641	ACATTGTACTATGAGTATACCAATTTTATTAAGAAATTTACATATTTTGATATCTTGATGT	2700
Oy	2701	ATTGTCTCTTGTCATTTACACAGATTTACTGTGTGTTTTTGTACTATCAAAATGTTCAAG	2760
Dp	2701	ATTGTCTCTTGTCATTTACACAGATTTACTGTGTGTTTTTGTACTATCAAAATGTTCAAG	2760
Oy	2761	GCAGCTCTTCCCTCAGATTTCTTGAGACACAAAGGAAGATTTCTTTAAGCGCATTTGGT	2820
Dp	2761	GCAGCTCTTCCCTCAGATTTCTTGAGACACAAAGGAAGATTTCTTTAAGCGCATTTGGT	2820
Oy	2821	CTGCTAAAGGAATCATCAAGATATGCTACAAACAATTAAGAAAACAATTAACATTAACA	2880
Dp	2821	CTGCTAAAGGAATCATCAAGATATGCTACAAACAATTAAGAAAACAATTAACATTAACA	2880
Oy	2881	TGTCCTCAACAGCCGAAGGAATGCAATGTTGTCATGTAACCTAATTTTGGAAAGTAAA	2940
Dp	2881	TGTCCTCAACAGCCGAAGGAATGCAATGTTGTCATGTAACCTAATTTTGGAAAGTAAA	2940
Oy	2941	AAATCTTGGAGAGTGCAGTATCATCAATAACTTATATAGATTAATCTGGAGCCGA	3000
Dp	2941	AAATCTTGGAGAGTGCAGTATCATCAATAACTTATATAGATTAATCTGGAGCCGA	3000
Oy	3001	ATGCATCCAAATTAATCTTCAAAATTCAAATTAACATTTCTTCGTACATATTTT	3060
Dp	3001	ATGCATCCAAATTAATCTTCAAAATTCAAATTAACATTTCTTCGTACATATTTT	3060
Oy	3061	TGAAGATCATGATTTTAAAGAAATAAGACGAGCTAAAGTTATGCTACGACATTAATCAT	3120
Dp	3061	TGAAGATCATGATTTTAAAGAAATAAGACGAGCTAAAGTTATGCTACGACATTAATCAT	3120
Oy	3121	CTGATATCTTTTGTGCATCTTTTGTATATCTGTGAATGTTAATGGTCAAAATCATATTT	3180
Dp	3121	CTGATATCTTTTGTGCATCTTTTGTATATCTGTGAATGTTAATGGTCAAAATCATATTT	3180
Oy	3181	ACACAAATATCCAGTCTAGTTTCTAAGAAAGATGATTTTCTGTCGATCAATCAATCC	3240
Dp	3181	ACACAAATATCCAGTCTAGTTTCTAAGAAAGATGATTTTCTGTCGATCAATCAATCC	3240
Oy	3241	GTATTAACTTCATGTAACAGGTGAACCTGAACCTTAATCTTTTGGAGAAATTAAGACA	3300
Dp	3241	GTATTAACTTCATGTAACAGGTGAACCTGAACCTTAATCTTTTGGAGAAATTAAGACA	3300
Oy	3301	TGACATTAATTTTGTCTGCAAGCTCGCAAAAGAAATCAGTAATCTTATGCCACAGGTAG	3360
Dp	3301	TGACATTAATTTTGTCTGCAAGCTCGCAAAAGAAATCAGTAATCTTATGCCACAGGTAG	3360

QY	3361	GAATTCATGTTGATTTTTTGACGTGATATGAAAGTTCTTATCAATTTCCGAGATGACATATA	3420
Db	3361	GAATTCATGTTGATTTTTTGACGTGATATGAAAGTTCTTATCAATTTCCGAGATGACATATA	3420
QY	3421	CATATTAATGATTCACATATATATGATGAGAAATGTGATTAACAAGTGTAGATATATCTGTG	3480
Db	3421	CATATTAATGATTCACATATATATGATGAGAAATGTGATTAACAAGTGTAGATATATCTGTG	3480
QY	3481	AAGACTTTTTTAACAACAATATTCGTGTGAAGCTAGATATCACTGATCTTCTCTGTTTC	3540
Db	3481	AAGACTTTTTTAAACAACAATATTCGTGTGAAGCTAGATATCACTGATCTTCTCTGTTTC	3540
QY	3541	TTGACGTGATCCCTGGTCACATGTTGTGTCCTCACAAAAAAATAGCAAGACATGTT	3600
Db	3541	TTGACGTGATCCCTGGTCACATGTTGTGTCCTCACAAAAAAATAGCAAGACATGTT	3600
QY	3601	TCAATAATATGTTAATAATAATAATTAGCCTTTAATTATATGCTTCTATTTGAGATAT	3660
Db	3601	TCAATAATATGTTAATAATAATAATTAGCCTTTAATTATATGCTTCTATTTGAGATAT	3660
QY	3661	TTTGTGATGCCACTTATATATTTGTGACATATCTCAAAAAAACAACCTTATATATGATG	3720
Db	3661	TTTGTGATGCCACTTATATATTTGTGACATATCTCAAAAAAACAACCTTATATATGATG	3720
QY	3721	CCTCTCAAAATGTAAGGAGTGTTCCTGGAATGGCAAACTGGGTCCGCAATACCTTTGCTTG	3780
Db	3721	CCTCTCAAAATGTAAGGAGTGTTCCTGGAATGGCAAACTGGGTCCGCAATACCTTTGCTTG	3780
QY	3781	TGTTCCATCTTCTCTTCAAAATGGTCTCCGAAAGATCAATCATTTCTGTCAAGGAACA	3840
Db	3781	TGTTCCATCTTCTCTTCAAAATGGTCTCCGAAAGATCAATCATTTCTGTCAAGGAACA	3840
QY	3841	GAAAGAAATTCGAGCGATGATTCCTAGTGTATCTGACTGAAGCTGTAAATCATCTCC	3900
Db	3841	GAAAGAAATTCGAGCGATGATTCCTAGTGTATCTGACTGAAGCTGTAAATCATCTCC	3900
QY	3901	CAGTATCCCACTGATATCTTTCATATAAATGGAACCTTTAGTCTCTATGAAATGAAAT	3960
Db	3901	CAGTATCCCACTGATATCTTTCATATAAATGGAACCTTTAGTCTCTATGAAATGAAAT	3960
QY	3961	CAACATCCCTGTAATATGTTCTGGGTGTTGGCCCGGAACGAACATAGGAATGTTAT	4020
Db	3961	CAACATCCCTGTAATATGTTCTGGGTGTTGGCCCGGAACGAACATAGGAATGTTAT	4020
QY	4021	GTTATGTAAGTTACATTGGCGCTCGAAGATCTTTGAAGTTTTTTTTTTTTTGGGGGGGG	4080
Db	4021	GTTATGTAAGTTACATTGGCGCTCGAAGATCTTTGAAGTTTTTTTTTTTTTGGGGGGGG	4080
QY	4081	GGGGGGGGGGGTGCTTTGATATTACTCTTAAAGTACAAGTTCTCTAAGTTATGTCAAGCA	4140
Db	4081	GGGGGGGGGGGTGCTTTGATATTACTCTTAAAGTACAAGTTCTCTAAGTTATGTCAAGCA	4140
QY	4141	CTTTGTAAACAATGTGATTTGGTATCATGATATGAGTTAACTAGTCAGATCTTGGT	4200
Db	4141	CTTTGTAAACAATGTGATTTGGTATCATGATATGAGTTAACTAGTCAGATCTTGGT	4200
QY	4201	AAGCAAAACCCTATCTATGTTAGGCTCACTAAAGTGGCGTTTGGTTCGAGAGAGAGAA	4260
Db	4201	AAGCAAAACCCTATCTATGTTAGGCTCACTAAAGTGGCGTTTGGTTCGAGAGAGAGAA	4260
QY	4261	GGATCAGTATGATATCCCAATCATCGAAGTAAATCATGTTGTTGGTACACATTTT	4320
Db	4261	GGATCAGTATGATATCCCAATCATCGAAGTAAATCATGTTGTTGGTACACATTTT	4320
QY	4321	CTACAAATCTAGTAGCTGATCGGTGAGCTACTGATCAACACCACTGCAACAACATATT	4380
Db	4321	CTACAAATCTAGTAGCTGATCGGTGAGCTACTGATCAACACCACTGCAACAACATATT	4380
QY	4381	CTCTGTGCAAAATGGGCAACCAAGATTAATCATCAAGCTGGAAGCAACCAAAATTTG	4440
Db	4381	CTCTGTGCAAAATGGGCAACCAAGATTAATCATCAAGCTGGAAGCAACCAAAATTTG	4440
QY	4441	AAGAGAGAAACCTTCACAAAGACTTTGAGTGGCCCCCACAAATCAGTTAGGCGCCG	4500

```

Db 4441 AAGAGAGAACCCCTCACAAACCTTGAAGTCCCGCCCAATCATGATGAGCCGCG 4500
Qy 4501 TCGAGGCGCGAGTGTACCATGCGACCAACCAATCCGAGGAGGACGTCAC 4560
Db 4501 TCGAGGCGCGAGTGTACCATGCGACCAACCAATCCGAGGAGGACGTCAC 4560
Qy 4561 CGATTACTGAATTCGCCCAACATCTTAATTTGTGAACAAATTTAAACAGAAACA 4620
Db 4561 CGATTACTGAATTCGCCCAACATCTTAATTTGTGAACAAATTTAAACAGAAACA 4620
Qy 4621 ATTTTGAATTTGTGAACAAATTTTAAACGGTATTCCTGAACATTTTCAAATTTGT 4680
Db 4621 ATTTTGAATTTGTGAACAAATTTTAAACGGTATTCCTGAACATTTTCAAATTTGT 4680
Qy 4681 GATCAAAATTTTAAACGACTCTCTTCTCAATTTGACATATTTTAAATTAATAAAA 4740
Db 4681 GATCAAAATTTTAAACGACTCTCTTCTCAATTTGACATATTTTAAATTAATAAAA 4740
Qy 4741 GTTCAACAAATTTTGAACATTTTAAACATTTAGCGAACATTTTGAATTTCTAATATTTT 4800
Db 4741 GTTCAACAAATTTTGAACATTTTAAACATTTAGCGAACATTTTGAATTTCTAATATTTT 4800
Qy 4801 CGAATTTGGAACATTTTCTAATTTCTGAACAAATTTGAAATACGAACGTAAATTTGA 4860
Db 4801 CGAATTTGGAACATTTTCTAATTTCTGAACAAATTTGAAATACGAACGTAAATTTGA 4860
Qy 4861 ATAAATTTTGGAAATGCGATTTTGTGAATTTCTGAACATATTTTGAATTAATAAAA 4920
Db 4861 ATAAATTTTGGAAATGCGATTTTGTGAATTTCTGAACATATTTTGAATTAATAAAA 4920
Qy 4921 CTTTAAAGGTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4980
Db 4921 CTTTAAAGGTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4980
Qy 4981 AAANAAATCCGAGAAAGCCCACTGGGAATGCAATGGAACAAACCCAGCCGTCCGCGC 5040
Db 4981 AAANAAATCCGAGAAAGCCCACTGGGAATGCAATGGAACAAACCCAGCCGTCCGCGC 5040
Qy 5041 ACTGTGAAAGCTAATTAAGTAGCGGCGCAAGCCCTCGTCTCATCATACCTGTGCGA 5100
Db 5041 ACTGTGAAAGCTAATTAAGTAGCGGCGCAAGCCCTCGTCTCATCATACCTGTGCGA 5100
Qy 5101 AACCCGACAAATTCGTGCACTATGCGCGAAATAGGCTTTTCCAGAGCTCTGTCTCC 5160
Db 5101 AACCCGACAAATTCGTGCACTATGCGCGAAATAGGCTTTTCCAGAGCTCTGTCTCC 5160
Qy 5161 GGTATAGGTCATTTTGAACCCCTCTCCACTGGGCGAGGCTAATTAATCTTTTTC 5220
Db 5161 GGTATAGGTCATTTTGAACCCCTCTCCACTGGGCGAGGCTAATTAATCTTTTTC 5220
Qy 5221 TTCTTTGCACTCACTTACTACGCGCAGTTTATGTTTGAAGGACCAACCGGTTTGT 5280
Db 5221 TTCTTTGCACTCACTTACTACGCGCAGTTTATGTTTGAAGGACCAACCGGTTTGT 5280
Qy 5281 GAAGGTCATGAACCTCAACCATTTTGGGAAGCTTCTAGAAGCCTATGAATGTTCTTT 5340
Db 5281 GAAGGTCATGAACCTCAACCATTTTGGGAAGCTTCTAGAAGCCTATGAATGTTCTTT 5340
Qy 5341 TGACATGATTAATTTGTTTCTTTTCAATTTGACAACTTTTTCAAATTCAT 5400
Db 5341 TGACATGATTAATTTGTTTCTTTTCAATTTGACAACTTTTTCAAATTCAT 5400
Qy 5401 GATTTTGTGAACCTTGTGAATTTTGAATCCGATTTTTCCTAATCCGTGTTT 5460
Db 5401 GATTTTGTGAACCTTGTGAATTTTGAATCCGATTTTTCCTAATCCGTGTTT 5460
Qy 5461 GAAAAAACTGTGACCTTTCCGAATTAATGAACATTTGTTGCAAGATGATGATCT 5520
Db 5461 GAAAAAACTGTGACCTTTCCGAATTAATGAACATTTGTTGCAAGATGATGATCT 5520
Qy 5521 TTTCAATGAGCGATTTTCTAATAATCCATATTTTCAATATTAAGCTTTCC 5580

```

```

Db 5521 TTTCAATGAGCGATTTTCTAATAATCCATATTTTCAATATTAAGCTTTCC 5580
Qy 5581 TTTTAAATCGAAGTATCTAGATTTGGAAGTCTTTAATTTCTTAATAAATGA 5640
Db 5581 TTTTAAATCGAAGTATCTAGATTTGGAAGTCTTTAATTTCTTAATAAATGA 5640
Qy 5641 TTTTAAATCGAAGTATCTAGATTTGGAAGTCTTTAATTTCTTAATAAATGA 5700
Db 5641 TTTTAAATCGAAGTATCTAGATTTGGAAGTCTTTAATTTCTTAATAAATGA 5700
Qy 5701 AACCATGATCTCGTGAACGCGGAGGCTTTCAATGATTTGGTGTCTAATACGCG 5760
Db 5701 AACCATGATCTCGTGAACGCGGAGGCTTTCAATGATTTGGTGTCTAATACGCG 5760
Qy 5761 CCCCTGACTTAACGAGAAATTTGAATTTTCAACACACTATGATGTAATAAATG 5820
Db 5761 CCCCTGACTTAACGAGAAATTTGAATTTTCAACACACTATGATGTAATAAATG 5820
Qy 5821 AATATCAGATTAATAAAGGGAATAAATAAATAAATAAATAAATAAATAAATAAATA 5880
Db 5821 AATATCAGATTAATAAAGGGAATAAATAAATAAATAAATAAATAAATAAATAAATA 5880
Qy 5881 ATTTGTTGTTGGGAGAACTAATCTGAAGAACTAATCTAAGTTCTCCGAATAAATA 5940
Db 5881 ATTTGTTGTTGGGAGAACTAATCTGAAGAACTAATCTAAGTTCTCCGAATAAATA 5940
Qy 5941 TTATGATTCGGGGAGATATTAATTAACAGATCAACCAAGATGTCCTCCAGAA 6000
Db 5941 TTATGATTCGGGGAGATATTAATTAACAGATCAACCAAGATGTCCTCCAGAA 6000
Qy 6001 ACCCAAGAAAGTCTAGAGTTTGTTCAGAGACCGAAAGATAGATAGCAATCTGA 6060
Db 6001 ACCCAAGAAAGTCTAGAGTTTGTTCAGAGACCGAAAGATAGATAGCAATCTGA 6060
Qy 6061 ATTAAGTCCATCTTTTCCAGAGATTTGAAGAAAGATAGATTTGAAATAGTGGC 6120
Db 6061 ATTAAGTCCATCTTTTCCAGAGATTTGAAGAAAGATAGATTTGAAATAGTGGC 6120
Qy 6121 GAGATATCATTTCTGATTAATGATTAATTTTGGCGGACAGCCAAACCCGCAATGG 6180
Db 6121 GAGATATCATTTCTGATTAATGATTAATTTTGGCGGACAGCCAAACCCGCAATGG 6180
Qy 6181 AGCCGGAATTTGAATTAAGTGGGTGAGAGTGAAGATGTTTCAATGAG 6240
Db 6181 AGCCGGAATTTGAATTAAGTGGGTGAGAGTGAAGATGTTTCAATGAG 6240
Qy 6241 TGTAGCAGATGATGATGCTCTTGTGTGAGCTGCAATATACATTAATAAATA 6300
Db 6241 TGTAGCAGATGATGATGCTCTTGTGTGAGCTGCAATATACATTAATAAATA 6300
Qy 6301 GAGCATCAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6360
Db 6301 GAGCATCAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6360
Qy 6361 TTGATGTGCGCAGGCGAGTGGCGACGCGAATTAAGCAACGCGCAACGCGC 6420
Db 6361 TTGATGTGCGCAGGCGAGTGGCGACGCGAATTAAGCAACGCGCAACGCGC 6420
Qy 6421 CGCGCTTCTCAATCCATCCATCCCTTAAGCTATTAATGAATGAATGATGTTCCCTG 6480
Db 6421 CGCGCTTCTCAATCCATCCATCCCTTAAGCTATTAATGAATGAATGATGTTCCCTG 6480
Qy 6481 CCACTGCTAGTACTCTCTCTGTTCTCTGTTGGAATGTTACACAGAGCAACGCGCA 6540
Db 6481 CCACTGCTAGTACTCTCTCTGTTCTCTGTTGGAATGTTACACAGAGCAACGCGCA 6540
Qy 6541 CGCGAGGCGCGCGCGCGCGCGCGCAACGCGCAACGCGCAACGCGCGCGCGGAA 6600
Db 6541 CGCGAGGCGCGCGCGCGCGCGCGCGCAACGCGCAACGCGCAACGCGCGCGCGGAA 6600
Qy 6601 CGGCAAGAGCAACGCGGACGCGGCGCGGCGGCGGAGTGAATTTTCCCGCGGCGCA 6660
Db 6601 CGGCAAGAGCAACGCGGACGCGGCGCGGCGGCGGAGTGAATTTTCCCGCGGCGCA 6660

```

6661 GGACGCGCATCTGCGGCGAGCGGGGCGAAGAACAGATCCGGGCGATACGGTACAAAT 6720
 6662 GGACGCGCATCTGCGGCGAGCGGGGCGAAGAACAGATCCGGGCGATACGGTACAAAT 6720
 6721 CAGCGCGAGCGTGAAGAGAGCGGGGCGCGCGCGTGTGCGCGTCCGCGCGAGCC 6780
 6722 CAGCGCGAGCGTGAAGAGAGCGGGGCGCGCGCGTGTGCGCGTCCGCGCGAGCC 6780
 6781 GTCCGCTGTTCCTCCGCGCTCCGCGAGCGCGTGTGCGCGAGCCGCGCGCT 6840
 6782 GTCCGCTGTTCCTCCGCGCTCCGCGAGCGCGTGTGCGCGAGCCGCGCGCT 6840
 6811 GCGCAGCGGCGAGTTCACCTGCTACGCGCGCGTGTGCGCGAGCCGCGCGAGGTA 6900
 6812 GCGCAGCGGCGAGTTCACCTGCTACGCGCGCGTGTGCGCGAGCCGCGCGAGGTA 6900
 6901 ACATTACAGCTTCACCGTAAATGTATGCGTGAAGCATGATGCGCGTTTACTTACGTC 6960
 6902 ACATTACAGCTTCACCGTAAATGTATGCGTGAAGCATGATGCGCGTTTACTTACGTC 6960
 6961 CCGCGCGTGTTCCTCCGCGCTCCGCGAGCGCGTGTGCGCGAGCCGCGCGAGGTA 7020
 6962 CCGCGCGTGTTCCTCCGCGCTCCGCGAGCGCGTGTGCGCGAGCCGCGCGAGGTA 7020
 7021 ACAAACAGCGCGTGAAGAGAGCTTCACGCGCGCGTGTGCGCGAGCCGCGCGAGGTA 7080
 7022 ACAAACAGCGCGTGAAGAGAGCTTCACGCGCGCGTGTGCGCGAGCCGCGCGAGGTA 7080
 7081 GACGCTTCCTCAACCGCGCGGAGCATGAGCGATCGAAGTCAATATCCGCGTGTGCGC 7140
 7082 GACGCTTCCTCAACCGCGCGGAGCATGAGCGATCGAAGTCAATATCCGCGTGTGCGC 7140
 7141 CAGACTGCGCGCGCGCAATATGCTTCCCGCGCGAGCTATCCAAATTAAGAGCGCGA 7200
 7142 CAGACTGCGCGCGCGCAATATGCTTCCCGCGCGAGCTATCCAAATTAAGAGCGCGA 7200
 7201 GCGGCGATTCACAGAGCTGAGAGTCCGCGAGCTTCGAGCTTCACCGCGAGGAGGAG 7260
 7202 GCGGCGATTCACAGAGCTGAGAGTCCGCGAGCTTCGAGCTTCACCGCGAGGAGGAG 7260
 7261 ATCGACATCGATCGCTGGAATCCATCGCGCGAGCAAGAACACCGCGAGTGTATCAT 7320
 7262 ATCGACATCGATCGCTGGAATCCATCGCGCGAGCAAGAACACCGCGAGTGTATCAT 7320
 7321 AACCGAACATCCGCTGCGCGAGCGTTCATCTGAGCATCTGCGCGAGGTTTGCAT 7380
 7322 AACCGAACATCCGCTGCGCGAGCGTTCATCTGAGCATCTGCGCGAGGTTTGCAT 7380
 7381 CCATGACATCTCTGCTGCTGATGAGACCGGCTGTGGAACATATGATGATGAGT 7440
 7382 CCATGACATCTCTGCTGCTGATGAGACCGGCTGTGGAACATATGATGATGAGT 7440
 7441 TTGCTAATCGTGTGCTGATGATGCTGTTGTTATCAGGTGCGCGAGTGTGCAAGAGC 7500
 7442 TTGCTAATCGTGTGCTGATGATGCTGTTGTTATCAGGTGCGCGAGTGTGCAAGAGC 7500
 7501 TCGGAATATGCTGATGCTGATGAGAGTTTACGCGCAACTGCTTGTGCGAGCGCGCT 7560
 7502 TCGGAATATGCTGATGCTGATGAGAGTTTACGCGCAACTGCTTGTGCGAGCGCGCT 7560
 7561 TTATCCGATGAGGCGCTTTTGGGCACTTGGCGCGCTGTGCTGATGATGATGATGAT 7620
 7562 TTATCCGATGAGGCGCTTTTGGGCACTTGGCGCGCTGTGCTGATGATGATGATGAT 7620
 7621 AGTCGTGATGATGCTGATGAGAGCTTGAATGAGTGTGCGAGTGTGCAAGAGC 7680
 7622 AGTCGTGATGATGCTGATGAGAGCTTGAATGAGTGTGCGAGTGTGCAAGAGC 7680
 7681 TTTTAGAAGAACTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 7740
 7682 TTTTAGAAGAACTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 7740

7741 GATTAGATTTTGTCTAAATTTGTACTGCTTGTGTTATTCAGATCTTACGCTATATAC 7800
 7742 GATTAGATTTTGTCTAAATTTGTACTGCTTGTGTTATTCAGATCTTACGCTATATAC 7800
 7801 GAATTAACCTTAATGCTCAACGAGAACCGAACCTTGTGCTGATGATGATGATGAT 7860
 7802 GAATTAACCTTAATGCTCAACGAGAACCGAACCTTGTGCTGATGATGATGATGAT 7860
 7861 GCGCTATTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7920
 7862 GCGCTATTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7920
 7921 AGAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7980
 7922 AGAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7980
 7981 TTTAGTATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8040
 7982 TTTAGTATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8040
 8041 AACCTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 8100
 8042 AACCTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 8100
 8101 TCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8160
 8102 TCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8160
 8161 GTTAAGAAATTTCTTACCTAACAACCTTGTGCTGATCTTCAATTTTGAATCTTGA 8220
 8162 GTTAAGAAATTTCTTACCTAACAACCTTGTGCTGATCTTCAATTTTGAATCTTGA 8220
 8221 CATTAATGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280
 8222 CATTAATGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280
 8281 TTAGGAAGCTCTTCTTAAATTTCTTGAAGAACCAAAAGCATTTCTTAAAGAGATT 8340
 8282 TTAGGAAGCTCTTCTTAAATTTCTTGAAGAACCAAAAGCATTTCTTAAAGAGATT 8340
 8341 TGTGCTAATAAGAAATCATGAGATGATGATGATGATGATGATGATGATGATGAT 8400
 8342 TGTGCTAATAAGAAATCATGAGATGATGATGATGATGATGATGATGATGATGAT 8400
 8401 TAGGCTCTCAAGAGCAAGAGAGATGATGATGATGATGATGATGATGATGATGAT 8460
 8402 TAGGCTCTCAAGAGCAAGAGAGATGATGATGATGATGATGATGATGATGATGAT 8460
 8461 CTTTAAAGGTTAATGAGATCTCAAGTCAATCAACCAATCAATCAATCAATCAAT 8520
 8462 CTTTAAAGGTTAATGAGATCTCAAGTCAATCAACCAATCAATCAATCAATCAAT 8520
 8521 TATGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8580
 8522 TATGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8580
 8581 AAATTTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8640
 8582 AAATTTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8640
 8641 ATTGTAACACTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 8700
 8642 ATTGTAACACTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 8700
 8701 CTTTCTAATAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8760
 8702 CTTTCTAATAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8760
 8761 GGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8820
 8762 GGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8820
 8821 TTTATGCTCAGGTAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 8880

Db 8821 TTTATGTCAGAGTGAATGATATAGGCCATTTTAAAGAAACATATGATATATAT 8880
Qy ATCTTCTGTATATCTAAACAATCTTCTCCATCCCTAAATGAATGCTTACCTTACG 8940
Db 8881 ATCTTCTGTATATCTAAACAATCTTCTCCATCCCTAAATGAATGCTTACCTTACG 8940
Qy ACAATTTTATATCTAGATCTAGTACCAAAAGTGAACAGTTATTTTGGACAGAGGAGTAG 9000
Db 8941 ACAATTTTATATCTAGATCTAGTACCAAAAGTGAACAGTTATTTTGGACAGAGGAGTAG 9000
Qy TATATATGTTGTGAGACATTAAGGTTATGTTGACCTGATATATGCTTCTTAAATGTGAAA 9060
Db 9001 TATATATGTTGTGAGACATTAAGGTTATGTTGACCTGATATATGCTTCTTAAATGTGAAA 9060
Qy CATGTTCTTATGTTTATGTTTATGTTATGTTATGAGAGTCTTACATTTCCGAGATGACTAC 9120
Db 9061 CATGTTCTTATGTTTATGTTTATGTTATGTTATGAGAGTCTTACATTTCCGAGATGACTAC 9120
Qy ACATTAATGATTACCATATCATTTGTGAGAAATGATATACCATTAAGATATCTTCTTCT 9180
Db 9121 ACATTAATGATTACCATATCATTTGTGAGAAATGATATACCATTAAGATATCTTCTTCT 9180
Qy TTTTATGCAAAAGACTAGCAGTGGCATGTATCTTTCTCTGTACCTATGCTCTTTTCTTTC 9240
Db 9181 TTTTATGCAAAAGACTAGCAGTGGCATGTATCTTTCTCTGTACCTATGCTCTTTTCTTTC 9240
Qy TCGTTACATGTTTGGCTCTCACAAAATAATATATCCAGACATGTTCCCAATATAT 9300
Db 9241 TCGTTACATGTTTGGCTCTCACAAAATAATATATCCAGACATGTTCCCAATATAT 9300
Qy ATTATATATTTTGAAGGTTTTCACCAACCTATATATCTTCTTCTTCTTCTTCTTCTTCT 9360
Db 9301 ATTATATATTTTGAAGGTTTTCACCAACCTATATATCTTCTTCTTCTTCTTCTTCTTCT 9360
Qy GTATATATGTTTAACTCTACCAAAACCTTATATATGTTTCTCTTCTTCTTCTTCTTCTTCT 9420
Db 9361 GTATATATGTTTAACTCTACCAAAACCTTATATATGTTTCTCTTCTTCTTCTTCTTCTTCT 9420
Qy TCTTGAATGGAAGAAATGGGTCGGTAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9480
Db 9421 TCTTGAATGGAAGAAATGGGTCGGTAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9480
Qy TGAAGCTGGAAGGCTCAAAATCATTTCTGTCAAAAGGAACAAGAAAGATCTTATATATG 9540
Db 9481 TGAAGCTGGAAGGCTCAAAATCATTTCTGTCAAAAGGAACAAGAAAGATCTTATATATG 9540
Qy TTTGTTAGTTGTACACACCCCTAGTTGTATCATCTGACGAGCTGTAAATCATTTCTAGTT 9600
Db 9541 TTTGTTAGTTGTACACACCCCTAGTTGTATCATCTGACGAGCTGTAAATCATTTCTAGTT 9600
Qy ATCCCACTTATATATTTCAATTAATAACATATGTTATGTTCTGTGAGTGTCAAGT 9660
Db 9601 ATCCCACTTATATATTTCAATTAATAACATATGTTATGTTCTGTGAGTGTCAAGT 9660
Qy CATGTTCTTATGTTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATG 9720
Db 9661 CATGTTCTTATGTTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATG 9720
Qy GGTGTTGTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9780
Db 9721 GGTGTTGTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9780
Qy CTATATCTTTCAACAATGAACCTTGAAGTCCCTTGAAGTGAAGTCAATATCTTCTTCTTCTTCT 9840
Db 9781 CTATATCTTTCAACAATGAACCTTGAAGTCCCTTGAAGTGAAGTCAATATCTTCTTCTTCTTCT 9840
Qy GAATATGCTATGTTTCAATGCTGTGATGATAATGATGATAATGATGATAATGATGATAATGATG 9900
Db 9841 GAATATGCTATGTTTCAATGCTGTGATGATAATGATGATAATGATGATAATGATGATAATGATG 9900
Qy CATGTTCAAGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG 9960
Db 9901 CATGTTCAAGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG 9960

Db 9901 CATGTTCAAGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG 9960
Qy ACATTTCTGTAAGTAAATTTGTCAGTTTGTATCTTGAAGGACCTTATATTTTCTTATTTGGG 10020
Db 9961 ACATTTCTGTAAGTAAATTTGTCAGTTTGTATCTTGAAGGACCTTATATTTTCTTATTTGGG 10020
Qy GGGGGGGGGTGAATAGCGGTTTATTAACAATTTGTTATTTGGAATATCTTAAATGTGGA 10080
Db 10021 GGGGGGGGGTGAATAGCGGTTTATTAACAATTTGTTATTTGGAATATCTTAAATGTGGA 10080
Qy ATTTAACTAGTGAATATTTTTTCCATTAAGGGTCTTTTATTTGACTCAATATGTAACA 10140
Db 10081 ATTTAACTAGTGAATATTTTTTCCATTAAGGGTCTTTTATTTGACTCAATATGTAACA 10140
Qy TCAAGGATACATATCAATTAAGTATGACATATGACATATCAATATAGTTGATGATCGG 10200
Db 10141 TCAAGGATACATATCAATTAAGTATGACATATGACATATCAATATAGTTGATGATCGG 10200
Qy CCAACACACACACACACACACACATTCACACACAAATCATGCTGACGAAGAGCGAA 10260
Db 10201 CCAACACACACACACACACACACATTCACACACAAATCATGCTGACGAAGAGCGAA 10260
Qy GTCATACAGATCAAAAATCTATGCTTGAAGGAGGAGGAATGAAAAATGAAAGAAATGA 10320
Db 10261 GTCATACAGATCAAAAATCTATGCTTGAAGGAGGAGGAATGAAAAATGAAAGAAATGA 10320
Qy AAAACGGTACATGACAAATCAATGACATGACGACGACAAATCATGCTGATGACAAACAAA 10380
Db 10321 AAAACGGTACATGACAAATCAATGACATGACGACGACAAATCATGCTGATGACAAACAAA 10380
Qy CTGCGAAGAAAGTTCTATTAATACTGGGCTTTCGAGAGAGGAAAGAGCTGCAAGATGTC 10440
Db 10381 CTGCGAAGAAAGTTCTATTAATACTGGGCTTTCGAGAGAGGAAAGAGCTGCAAGATGTC 10440
Qy CATCATCGATCCACACATTAAGTATCTGCGGTTTCTCATCTGAAATCAATCCG 10500
Db 10441 CATCATCGATCCACACATTAAGTATCTGCGGTTTCTCATCTGAAATCAATCCG 10500
Qy AGCAAACTCCGATTAATGTTTATTAAGGTTACGATTCGAAATGCGCAATCATGAG 10560
Db 10501 AGCAAACTCCGATTAATGTTTATTAAGGTTACGATTCGAAATGCGCAATCATGAG 10560
Qy TTAATGACAAATTAAGACACACCTAGATTTTATTCAAAGCTCGAGCGGTTACTTGA 10620
Db 10561 TTAATGACAAATTAAGACACACCTAGATTTTATTCAAAGCTCGAGCGGTTACTTGA 10620
Qy TTAATGACAAATTAAGACACACCTAGATTTTATTCAAAGCTCGAGCGGTTACTTGA 10680
Db 10621 TTAATGACAAATTAAGACACACCTAGATTTTATTCAAAGCTCGAGCGGTTACTTGA 10680
Qy AGTACCATCCATTTGAAGTATCTCCACTTGCCTCAATCAAAATGTTGCATATGTCAG 10680
Db 10681 AGTACCATCCATTTGAAGTATCTCCACTTGCCTCAATCAAAATGTTGCATATGTCAG 10680
Qy GTCCATATGCGAGTATGGAATGGAATGAGCGCGCATGTTAGTTTAAAGTGAAGTGAACA 10740
Db 10681 GTCCATATGCGAGTATGGAATGGAATGAGCGCGCATGTTAGTTTAAAGTGAAGTGAACA 10740
Qy GCCTGTGCGCCACCACTGAGGAAGTGTGATGAGGAGGAAGATATGCTCCACCAAC 10800
Db 10741 GCCTGTGCGCCACCACTGAGGAAGTGTGATGAGGAGGAAGATATGCTCCACCAAC 10800
Qy ATCCCAAGTTTGAACAATTTCTAGAGCCCTTACCAATCTCAACAAGCGAAATTTGATGAC 10860
Db 10801 ATCCCAAGTTTGAACAATTTCTAGAGCCCTTACCAATCTCAACAAGCGAAATTTGATGAC 10860
Qy TATCTGTATGACAGCAAAATCATATGCTGCTGCTATCTGCTATCTTGTGATGACATAC 10920
Db 10861 TATCTGTATGACAGCAAAATCATATGCTGCTGCTATCTGCTATCTTGTGATGACATAC 10920
Qy TACCTGCAAAAGGCGGATTCAGAGCTTCAGAGCTGCGGCGCGC 10966
Db 10921 TACCTGCAAAAGGCGGATTCAGAGCTTCAGAGCTGCGGCGCGC 10966
RESULT 2
BD012010

LOCUS	BD012010	10966 bp	DNA	linear	PAT 02-ANG-2002
DEFINITION	Creation of iron-deficiency resistant rice plant.				
ACCESSION	BD012010				
VERSION	BD012010.1	GI:22092199			
KEYWORDS	WO 0101762-A/L.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 10966)				
AUTHORS	Mori,S., Nakamichi,H., Takahashi,M. and Nishizawa,N.				
TITLE	Creation of iron-deficiency resistant rice plant				
JOURNAL	Patent: WO 0101762-A 1 11-JAN-2001;				
	JAPAN SCIENCE AND TECHNOLOGY CORP,SATOSHI MORI,HIROMI NAKANISHI,				
	MICHIO TAKAHASHI,NAOKO NISHIZAWA				
	OS Hordeum vulgare L. var. Igri				
	PN WO 0101762-A/1				
	PD 11-JAN-2001				
	PF 04-JUL-2000 WO 2000JP004425				
	PR 05-JUL-1999 JP 99P 190318				
	PI SATOSHI MORI,HIROMI NAKANISHI,MICHIO TAKAHASHI, PI NAOKO NISHIZAWA				
	PC A01H5/00,C12N5/14,C12N15/52				
	CC				
FEATURES	FH Key Location/Qualifiers.				
source	1..10966				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	100.0%;	Score 10966;	DB 6;	Length 10966;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 10966;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CTCGATCCCATTCGCAATGCTATGATTAGCTATCAACGAAAGAAAGAGATGCGATGCGC	60		
DB	1	CTCGATCCCATTCGCAATGCTATGATTAGCTATCAACGAAAGAAAGAGATGCGATGCGC	60		
QY	61	CTGTGTGTCATCCCTCACTGGCTTGGCGAATGGCGATACCGAGTTAGTAGAGTGTTTT	120		
DB	61	CTGTGTGTCATCCCTCACTGGCTTGGCGAATGGCGATACCGAGTTAGTAGAGTGTTTT	120		
QY	121	TTAGCATGATGTCTGCGCGCACTGCCAAGAAAATCGCGCTCAGCGCACTGCAGAGAGTT	180		
DB	121	TTAGCATGATGTCTGCGCGCACTGCCAAGAAAATCGCGCTCAGCGCACTGCAGAGAGTT	180		
QY	181	GAGCGATGATGCTTTGTGATGAGCGGAGCTGAGTGGGTCTCACTAACTGAACCAATCA	240		
DB	181	GAGCGATGATGCTTTGTGATGAGCGGAGCTGAGTGGGTCTCACTAACTGAACCAATCA	240		
QY	241	GCAATGGGTGATGAGTTCGACGAAGCATCATGCTTCTCGGCTCCCGATCCGCTTACTTT	300		
DB	241	GCAATGGGTGATGAGTTCGACGAAGCATCATGCTTCTCGGCTCCCGATCCGCTTACTTT	300		
QY	301	TTCTCCCAAAATTATTAAGAGAGATAGATGATGATGATGATGATGATGATGATGATGATG	360		
DB	301	TTCTCCCAAAATTATTAAGAGAGATAGATGATGATGATGATGATGATGATGATGATGATG	360		
QY	361	ATGAGAACCAAGCGAGGGCGCGAAATATGCGCGGATTAATGATGCGACCGCGCAACGCG	420		
DB	361	ATGAGAACCAAGCGAGGGCGCGAAATATGCGCGGATTAATGATGCGACCGCGCAACGCG	420		
QY	421	CACGCGCGCTCAGCTGCGCAGCGCGCTGCGCAGAGCTGCGCAGAGCGTGCAGCGCTGCT	480		
DB	421	CACGCGCGCTCAGCTGCGCAGCGCGCTGCGCAGAGCTGCGCAGAGCGTGCAGCGCTGCT	480		
QY	481	TCTTGCTCCGAGCGCGCGCTTCTGTGCGCGGTACAGCAACGAGCTATATAGAACCGTCAATC	540		
DB	481	TCTTGCTCCGAGCGCGCGCTTCTGTGCGCGGTACAGCAACGAGCTATATAGAACCGTCAATC	540		
QY	541	ACCGTACTCAATCCGTCGCCCAACTGCTTCTTAATACGCTACTAGTAGTATTCCTGGT	600		

[illegible]

QY 1681 GCATTTGCACTCATCCCGCAAGAGGAGTGGAGATGACATGACATGCTGATGATCAT 1740
DB 1681 GCATTTGCACTCATCCCGCAAGAGGAGTGGAGATGACATGACATGCTGATGATCAT 1740
QY 1741 CGCGGACAAAGAACACACCGGCGATGATCATCAAAACCCCAACAACCCGTCGGGAGCGT 1800
DB 1741 CGCGGACAAAGAACACACCGGCGATGATCAAAACCCCAACAACCCGTCGGGAGCGT 1800
QY 1801 TTACTCCAGACGACATGTCGCAAGGTTGACATCCTTTGGCTTGCTGGAATATGA TCA 1860
DB 1801 TTACTCCAGACGACATGTCGCAAGGTTGACATCCTTTGGCTTGCTGGAATATGA TCA 1860
QY 1861 GTTCAGTGCACCTGTCGTAATCTTTTGGCAATGCAATGCAATGATGATGCTCAATTA 1920
DB 1861 GTTCAGTGCACCTGTCGTAATCTTTTGGCAATGCAATGCAATGATGATGCTCAATTA 1920
QY 1921 GGTGCGGAGAGTGGCGAAAGAGCTCGAAATATGGTGATGCTGACGAGATACGGCAA 1980
DB 1921 GGTGCGGAGAGTGGCGAAAGAGCTCGAAATATGGTGATGCTGACGAGATACGGCAA 1980
QY 1981 GCTGATTCGAGGAGCGCCCGGTCATCCCAATGGAGTGTTTGGGCAATCACCCCTGT 2040
DB 1981 GCTGATTCGAGGAGCGCCCGGTCATCCCAATGGAGTGTTTGGGCAATCACCCCTGT 2040
QY 2041 GCTGTCAATAGGAGTCTGTGCAAGTCAATGATGATGCTGATGCGGCTTGATGGGT 2100
DB 2041 GCTGTCAATAGGAGTCTGTGCAAGTCAATGATGATGCTGATGCGGCTTGATGGGT 2100
QY 2101 AGCGGTGATGACACCCGAGAAAGATCTTACAGAAACATGATCTTAATCTCTAATCA 2160
DB 2101 AGCGGTGATGACACCCGAGAAAGATCTTACAGAAACATGATCTTAATCTCTAATCA 2160
QY 2161 TTCTTTTCAATGCTACTAAGGATTTAATAGTACTGATCAATATATTTGCTAAT 2220
DB 2161 TTCTTTTCAATGCTACTAAGGATTTAATAGTACTGATCAATATATTTGCTAAT 2220
QY 2221 TTGTACTGACATTTTGTGTAGATCTCTACATCAATTAAGCAATGCTGATGCTGCA 2280
DB 2221 TTGTACTGACATTTTGTGTAGATCTCTACATCAATTAAGCAATGCTGATGCTGCA 2280
QY 2281 CAGACCCAGACACCTTCAATGATGATCTTGTGATTTTACCTGCTTCAAGAAATTA 2340
DB 2281 CAGACCCAGACACCTTCAATGATGATCTTGTGATTTTACCTGCTTCAAGAAATTA 2340
QY 2341 GTCTTTGGATTTTACTCTCTCTGCTCAATTTGCTCGGCTCTTAAGTGTAGGCTGCC 2400
DB 2341 GTCTTTGGATTTTACTCTCTCTGCTCAATTTGCTCGGCTCTTAAGTGTAGGCTGCC 2400
QY 2401 CACGTGATGATGATGACGCTTTTTCATTAAGTTGAAAGTCAAAAGTCAAGACAT 2460
DB 2401 CACGTGATGATGATGACGCTTTTTCATTAAGTTGAAAGTCAAAAGTCAAGACAT 2460
QY 2461 ACATTTGATGATTTTATCTTTGCTTGTGATCCGATTAATTAATAAAATTAATAAA 2520
DB 2461 ACATTTGATGATTTTATCTTTGCTTGTGATCCGATTAATTAATAAAATTAATAAA 2520
QY 2521 CTGAACCTACTGTAATATATTAACACATGCTTCTTCAAGATATGATGATGCTATGGG 2580
DB 2521 CTGAACCTACTGTAATATATTAACACATGCTTCTTCAAGATATGATGATGCTATGGG 2580
QY 2581 CATGCCATATCTTTTGGGTCAGATGACATGATGATGAAACCTCTTTTGAATAATGAT 2640
DB 2581 CATGCCATATCTTTTGGGTCAGATGACATGATGATGAAACCTCTTTTGAATAATGAT 2640
QY 2641 ACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
DB 2641 ACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
QY 2701 ATGTTCTCTTGTGATTCACAGATTTTACTGTTGTTTGTGATGATGATGATGATGATGAT 2760
DB 2701 ATGTTCTCTTGTGATTCACAGATTTTACTGTTGTTTGTGATGATGATGATGATGATGAT 2760

QY 2761 GCAGCTCTTCTCAGATTTCTTGAGAACACAAAGAAATTTCTTTAAGCGATATTTGGT 2820
DB 2761 GCAGCTCTTCTCAGATTTCTTGAGAACACAAAGAAATTTCTTTAAGCGATATTTGGT 2820
QY 2821 CTGCTAAAGAAATCATCAGATATGCTACAAACAAATTAAGGAAACAAATCATTTACA 2880
DB 2821 CTGCTAAAGAAATCATCAGATATGCTACAAACAAATTAAGGAAACAAATCATTTACA 2880
QY 2881 TGTCTCACAAGGACAGAAAGATCAATGTTGTGATGATGATGATGATGATGATGATGAT 2940
DB 2881 TGTCTCACAAGGACAGAAAGATCAATGTTGTGATGATGATGATGATGATGATGATGAT 2940
QY 2941 AAATCTTGAAGAGTGCATGATATCATTAATTTATATGATATGATGATGATGATGATGAT 3000
DB 2941 AAATCTTGAAGAGTGCATGATATCATTAATTTATATGATATGATGATGATGATGATGAT 3000
QY 3001 ATGATTCACATTAATTTACTTCAATTCAAATTCAAATTCATTTCTTCCGATCATTTT 3060
DB 3001 ATGATTCACATTAATTTACTTCAATTCAAATTCAAATTCATTTCTTCCGATCATTTT 3060
QY 3061 TGAAGATGATGATTTTATTAAGATTAATGACGAGAGCTTAAATGATGATGATGATGAT 3120
DB 3061 TGAAGATGATGATTTTATTAAGATTAATGACGAGAGCTTAAATGATGATGATGATGAT 3120
QY 3121 CTGATATCTTGTGCTCATCTTTTGTATATCTGATGATGATGATGATGATGATGATGAT 3180
DB 3121 CTGATATCTTGTGCTCATCTTTTGTATATCTGATGATGATGATGATGATGATGATGAT 3180
QY 3181 ACAAAATATCCATGCTATGATTTCTAAGAAATGATGATGATGATGATGATGATGATGAT 3240
DB 3181 ACAAAATATCCATGCTATGATTTCTAAGAAATGATGATGATGATGATGATGATGATGAT 3240
QY 3241 GTATTAATCTTCCATGATTAACAGGTGAAACGAACTTATGATGATGATGATGATGATGAT 3300
DB 3241 GTATTAATCTTCCATGATTAACAGGTGAAACGAACTTATGATGATGATGATGATGATGAT 3300
QY 3301 TGAATGATTTTGTGCTGCAAGCTCGCAAAAGAAATGATGATGATGATGATGATGATGAT 3360
DB 3301 TGAATGATTTTGTGCTGCAAGCTCGCAAAAGAAATGATGATGATGATGATGATGATGAT 3360
QY 3361 GAATCCATGATTTTGTGCTGCAAGCTCGCAAAAGAAATGATGATGATGATGATGATGAT 3420
DB 3361 GAATCCATGATTTTGTGCTGCAAGCTCGCAAAAGAAATGATGATGATGATGATGATGAT 3420
QY 3421 CATATTAATGATTAACATATTTATGATGATGATGATGATGATGATGATGATGATGAT 3480
DB 3421 CATATTAATGATTAACATATTTATGATGATGATGATGATGATGATGATGATGATGAT 3480
QY 3481 AAGACTTTTAAACAAATTTCTGAGAGATGATGATGATGATGATGATGATGATGATGAT 3540
DB 3481 AAGACTTTTAAACAAATTTCTGAGAGATGATGATGATGATGATGATGATGATGATGAT 3540
QY 3541 TTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
DB 3541 TTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
QY 3601 TCAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660
DB 3601 TCAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660
QY 3661 TTTTGTAGTCAACCTATATATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3720
DB 3661 TTTTGTAGTCAACCTATATATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3720
QY 3721 CCTCTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
DB 3721 CCTCTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
QY 3781 TGTTCATCTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3840
DB 3781 TGTTCATCTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3840
QY 3841 GAAAGAAATTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900

```
Db 3841 GAAAGAAATTCGAGCGATGATTCCTAGTGTATATCTGACTGAGCTGAAATCAATTC 3900
Qy 3901 CAGTATCCCCCATATATCTTTCAATAAATGAACTTTAGTCTCTATGATGAAAGT 3960
Db 3901 CAGTATCCCCCATATATCTTTCAATAAATGAACTTTAGTCTCTATGATGAAAGT 3960
Qy 3961 CAACATCTCCCTGAAATATGTTCTGTTGTTGTCGCTGGAAGAAACATAGGAAATGTTAT 4020
Db 3961 CAACATCTCCCTGAAATATGTTCTGTTGTTGTCGCTGGAAGAAACATAGGAAATGTTAT 4020
Qy 4021 GTTAGTAAAGTTACATGGCGTCGAAAGATCTTTGAAGTTTTTTTTTTTTTTGGGGGGGG 4080
Db 4021 GTTAGTAAAGTTACATGGCGTCGAAAGATCTTTGAAGTTTTTTTTTTTTTTGGGGGGGG 4080
Qy 4081 GGGGGGGGGGCTTTGATATATCTTTAGTACAGTTCTCTCAAGTATGTCAAGCA 4140
Db 4081 GGGGGGGGGGCTTTGATATATCTTTAGTACAGTTCTCTCAAGTATGTCAAGCA 4140
Qy 4141 CTTTGTAAACAAATTTGATTTGATATGATATGATTAAGTAACTAGTCAATCTTGGT 4200
Db 4141 CTTTGTAAACAAATTTGATTTGATATGATATGATTAAGTAACTAGTCAATCTTGGT 4200
Qy 4201 AAGCAAAACCTTACTATGTTAGCTCACTAAGGTGCGTTTGGTTGAGAGAGAGAA 4260
Db 4201 AAGCAAAACCTTACTATGTTAGCTCACTAAGGTGCGTTTGGTTGAGAGAGAGAA 4260
Qy 4261 GGATCACTGATATATCCCAATCATGAAATGAAATCATGTTGTTGCTACCACTTTT 4320
Db 4261 GGATCACTGATATATCCCAATCATGAAATGAAATCATGTTGTTGCTACCACTTTT 4320
Qy 4321 CTACAACTCTAGTCTGAGCGTGAAGCTAGTACATGATACACACAGTCAACATAT 4380
Db 4321 CTACAACTCTAGTCTGAGCGTGAAGCTAGTACATGATACACACAGTCAACATAT 4380
Qy 4381 CTCTGTCAAAATCGGACCCCAAGATTAATCATCTCAAGCTGAAAGCAACCAATTTG 4440
Db 4381 CTCTGTCAAAATCGGACCCCAAGATTAATCATCTCAAGCTGAAAGCAACCAATTTG 4440
Qy 4441 AAGAGAGAACCTCTACAAAGACCTTTGATGTCCTCCCAACATGATGTTAGCCGCG 4500
Db 4441 AAGAGAGAACCTCTACAAAGACCTTTGATGTCCTCCCAACATGATGTTAGCCGCG 4500
Qy 4501 TCGAGGCGGAGTGCATGCGAGCAACCACTCCAGCGGGGAGACGTCAC 4560
Db 4501 TCGAGGCGGAGTGCATGCGAGCAACCACTCCAGCGGGGAGACGTCAC 4560
Qy 4561 CGATTACTGAAATTCCTCAACAAATCTTAATTTGTGAACAAATTTAAACAGAAACA 4620
Db 4561 CGATTACTGAAATTCCTCAACAAATCTTAATTTGTGAACAAATTTAAACAGAAACA 4620
Qy 4621 ATTTTGAATTTGTGAACAAATTTTAAACGSGTATCTGAAACATTTTCAAAATTTGT 4680
Db 4621 ATTTTGAATTTGTGAACAAATTTTAAACGSGTATCTGAAACATTTTCAAAATTTGT 4680
Qy 4681 GATCAAAATTTTAAAGACTCTTCTCAATTTGAGCAATTTTAAATTTAAATTTAA 4740
Db 4681 GATCAAAATTTTAAAGACTCTTCTCAATTTGAGCAATTTTAAATTTAAATTTAA 4740
Qy 4741 GTTCAACAAATTTTGAATTTTAAATTTAGCGAAACATTTGAAATTTCTAAATTTT 4800
Db 4741 GTTCAACAAATTTTGAATTTTAAATTTAGCGAAACATTTGAAATTTCTAAATTTT 4800
Qy 4801 CGAAATTTGGAACATTTTCTATTTCTGAACAAATTTGAAATTTGAAATTTGGA 4860
Db 4801 CGAAATTTGGAACATTTTCTATTTCTGAACAAATTTGAAATTTGAAATTTGGA 4860
Qy 4861 ATTAATTTTGGAAATGCGATTTTGTGAATTTCTGAACATTTTGAATTTGAAATTT 4920
Db 4861 ATTAATTTTGGAAATGCGATTTTGTGAATTTCTGAACATTTTGAATTTGAAATTT 4920
Qy 4921 CTTTAAAGGTAAATTAATAATTAATAATGAACATTAATAATGAACAAATTA 4980
Db 4921 CTTTAAAGGTAAATTAATAATTAATAATGAACATTAATAATGAACAAATTA 4980
```

```
Db 4921 CTTTAAAGGTAAATTAATAATTAATAATGAACATTAATAATGAACAAATTA 4980
Qy 4981 AAGAAATTCGAGAAACCAACTGGGAATGCAATGAAACCAAGCGCTCGCGCG 5040
Db 4981 AAGAAATTCGAGAAACCAACTGGGAATGCAATGAAACCAAGCGCTCGCGCG 5040
Qy 5041 ACTGTAAAGCTTAATGAGAGCGGCGCAAGCTCGTCTCATCATACCTGTGGA 5100
Db 5041 ACTGTAAAGCTTAATGAGAGCGGCGCAAGCTCGTCTCATCATACCTGTGGA 5100
Qy 5101 AACCCGACAAATTTGTTGACTATAGCGCGAATAGGCTTTTCCAGAGCTCTCTTCC 5160
Db 5101 AACCCGACAAATTTGTTGACTATAGCGCGAATAGGCTTTTCCAGAGCTCTCTTCC 5160
Qy 5161 GATTATGGGCTATTTGACACCCCTCTCCACTTGGGCGAGCTATATATTTTTC 5220
Db 5161 GATTATGGGCTATTTGACACCCCTCTCCACTTGGGCGAGCTATATATTTTTC 5220
Qy 5221 TTTCTTGACCTCACTGTTACTACGCGAGTTTGTGTTGGAAGGACCAACCGGTTTGT 5280
Db 5221 TTTCTTGACCTCACTGTTACTACGCGAGTTTGTGTTGGAAGGACCAACCGGTTTGT 5280
Qy 5281 GAAAGTTCTAGAACTCAACCAATTTTGGAGGCTCTAGAGGCTATGAAATGTTCTT 5340
Db 5281 GAAAGTTCTAGAACTCAACCAATTTTGGAGGCTCTAGAGGCTATGAAATGTTCTT 5340
Qy 5341 TGACATGATATATTTGTTTCTTTTCAATTTGACAAATCTTTTCAATTCAT 5400
Db 5341 TGACATGATATATTTGTTTCTTTTCAATTTGACAAATCTTTTCAATTCAT 5400
Qy 5401 GATTTTGTGAACTTTGATTTTGAATCCGATTTTTTTTCTAAATCCGTTT 5460
Db 5401 GATTTTGTGAACTTTGATTTTGAATCCGATTTTTTTTCTAAATCCGTTT 5460
Qy 5461 GAAAAAACTGTGACCTTTTCGAAATTAATGAACATTTGTTGCAAGATGATGATCT 5520
Db 5461 GAAAAAACTGTGACCTTTTCGAAATTAATGAACATTTGTTGCAAGATGATGATCT 5520
Qy 5521 TTTCAATGAGCGATTTTCTTAAATATCCAAATTTTCAATTTCAATTTTCC 5580
Db 5521 TTTCAATGAGCGATTTTCTTAAATATCCAAATTTTCAATTTTCAATTTTCC 5580
Qy 5581 TTTTAACTGAACTATCTTGAATTTGTTGTTGTTTAAATTTTCTTAAATTTTGA 5640
Db 5581 TTTTAACTGAACTATCTTGAATTTGTTGTTGTTTAAATTTTCTTAAATTTTGA 5640
Qy 5641 TTTTAACTGAACTATCTTGAATTTGTTGTTGTTTAAATTTTCTTAAATTTTGA 5700
Db 5641 TTTTAACTGAACTATCTTGAATTTGTTGTTGTTTAAATTTTCTTAAATTTTGA 5700
Qy 5701 AACCATTTGACTGCTGAACAGGGGCTTTCAATGATTTGGTGTCTTAATCCAGCG 5760
Db 5701 AACCATTTGACTGCTGAACAGGGGCTTTCAATGATTTGGTGTCTTAATCCAGCG 5760
Qy 5761 CCCCTGACTATAACGAAGAAATTTTGAACCACTAATTTTGAATTTTGAATTTG 5820
Db 5761 CCCCTGACTATAACGAAGAAATTTTGAACCACTAATTTTGAATTTTGAATTTG 5820
Qy 5821 AATATCAAGTAAATAAGGGAATAAATACTAATCCCTGAATTCCTGTCTTAAT 5880
Db 5821 AATATCAAGTAAATAAGGGAATAAATACTAATCCCTGAATTCCTGTCTTAAT 5880
Qy 5881 ATTTGTTGTTGGGAGAACTAATCTGAAGAACTAATCTGATCTCCGGAATAACAATA 5940
Db 5881 ATTTGTTGTTGGGAGAACTAATCTGAAGAACTAATCTGATCTCCGGAATAACAATA 5940
Qy 5941 TTATGATTCGGGGGAGTATTAATTAATCAAGATCAACCAAGATGCTCTCAAGAAA 6000
Db 5941 TTATGATTCGGGGGAGTATTAATTAATCAAGATCAACCAAGATGCTCTCAAGAAA 6000
Qy 6001 ACCCAAGAAAGTCTAGAGTTTGTCTTCAAGGACGAAAGATGATGATCTTGA 6060
Db 6001 ACCCAAGAAAGTCTAGAGTTTGTCTTCAAGGACGAAAGATGATGATCTTGA 6060
```

QY 6061 ATTAGTCCATCTTTTCCCAAGATTGAAAGAAAGATAGATTCTGAATTAGTGGC 6120
DB 6061 ATTAGTCCATCTTTTCCCAAGATTGAAAGAAAGATAGATTCTGAATTAGTGGC 6120
QY 6121 GAGATTCATTTCTGAGTATAGATCAATTTGTTTGGCCGACAGCCAAACCCCGAGTGG 6180
DB 6121 GAGATTCATTTCTGAGTATAGATCAATTTGTTTGGCCGACAGCCAAACCCCGAGTGG 6180
QY 6181 AGCCGGAATTGGAATTGAGTGGTGGAGTGGAGAGATGGTTCATGCGTTCTGAAGAG 6240
DB 6181 AGCCGGAATTGGAATTGAGTGGTGGAGTGGAGAGATGGTTCATGCGTTCTGAAGAG 6240
QY 6241 TGTAGCCAGTAGTGTGTCTCTTGTGTGTGAGTGCATATACAGATCAATTAACAA 6300
DB 6241 TGTAGCCAGTAGTGTGTCTCTTGTGTGTGAGTGCATATACAGATCAATTAACAA 6300
QY 6301 GACGATCAGCTGGACGGGTGCTGCTGATGCGTCTTCTGCGCCGCCAGAGCCCGG 6360
DB 6301 GACGATCAGCTGGACGGGTGCTGCTGATGCGTCTTCTGCGCCGCCAGAGCCCGG 6360
QY 6361 TTGATGTGCGCAGGCGAGTGGCGACGGGACCGAGCGCTATTAAGACAGGCGCAACGCG 6420
DB 6361 TTGATGTGCGCAGGCGAGTGGCGACGGGACCGAGCGCTATTAAGACAGGCGCAACGCG 6420
QY 6421 CGCGGTTCTCAATCCATCCATCCCTTAGCTGATTGATTGACTAGTTCATTTCCCTG 6480
DB 6421 CGCGGTTCTCAATCCATCCATCCCTTAGCTGATTGATTGACTAGTTCATTTCCCTG 6480
QY 6481 CCACATCGTCAATCT 6540
DB 6481 CCACATCGTCAATCT 6540
QY 6541 CGCGGAGGCG 6600
DB 6541 CGCGGAGGCG 6600
QY 6601 CGCGGAGGCG 6660
DB 6601 CGCGGAGGCG 6660
QY 6661 GGAAGGCAATCTGCG 6720
DB 6661 GGAAGGCAATCTGCG 6720
QY 6721 CACGCGCAGGCGTGGAGAGAGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6780
DB 6721 CACGCGCAGGCGTGGAGAGAGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6780
QY 6781 GTCGCGTGTCCCGGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 6840
DB 6781 GTCGCGTGTCCCGGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 6840
QY 6841 GCGCGCAGGCG 6900
DB 6841 GCGCGCAGGCG 6900
QY 6901 ACATTTACAGCTTCAACCGTATGATGCTGAGCATGCGCGCGGTTTACTTACGTGC 6960
DB 6901 ACATTTACAGCTTCAACCGTATGATGCTGAGCATGCGCGCGGTTTACTTACGTGC 6960
QY 6961 CGCGCGGCTGTTCTTCCCGGCTGCGTTCAAAATTTTAACTTTTAACTTTTAACTTT 7020
DB 6961 CGCGCGGCTGTTCTTCCCGGCTGCGTTCAAAATTTTAACTTTTAACTTTTAACTTT 7020
QY 7021 ACAAAAGGCGCGCGTGAAGAGCACTTGTCAAGGCGCTGCGCTTCAAGAGTATGCGCGGAC 7080
DB 7021 ACAAAAGGCGCGCGTGAAGAGCACTTGTCAAGGCGCTGCGCTTCAAGAGTATGCGCGGAC 7080
QY 7081 GAGCTTCTCTCAACCGCGCGCGAGACTCAGGCGATCGAATCAATATCCCGGTGCTGGC 7140
DB 7081 GAGCTTCTCTCAACCGCGCGCGAGACTCAGGCGATCGAATCAATATCCCGGTGCTGGC 7140

QY 7141 CAGACTGCGCGCGCGCAACTACTGCTTCCCGGCGCAGGCTATCCAAATTAAGAGCGCGA 7200
DB 7141 CAGACTGCGCGCGCGCAACTACTGCTTCCCGGCGCAGGCTATCCAAATTAAGAGCGCGA 7200
QY 7201 GCGGCAATTCAACAGCTGAGGCTCCGCGCATTCGACTCATCCCGCAAGAGGCTGGAG 7260
DB 7201 GCGGCAATTCAACAGCTGAGGCTCCGCGCATTCGACTCATCCCGCAAGAGGCTGGAG 7260
QY 7261 ATCGACATGACCTGCGTGGAAATCATGCGCGCAAGAAACCAACCGGAGTGCATATA 7320
DB 7261 ATCGACATGACCTGCGTGGAAATCATGCGCGCAAGAAACCAACCGGAGTGCATATA 7320
QY 7321 AACCAAAACATCCGTCGCGCAGCGCTTACTCTCAACACCATCTGCGCAAGGTTTTCAT 7380
DB 7321 AACCAAAACATCCGTCGCGCAGCGCTTACTCTCAACACCATCTGCGCAAGGTTTTCAT 7380
QY 7381 CCATGCACTCTCTGCTCTGCTGATCGACCGGCTGCTTTGAACATATGATGCTGCT 7440
DB 7381 CCATGCACTCTCTGCTCTGCTGATCGACCGGCTGCTTTGAACATATGATGCTGCT 7440
QY 7441 TTGCTATGCTGTGCTGATGATGCTGTTTGTATCAGGTCGCGGAGTGGCAAGAGAGC 7500
DB 7441 TTGCTATGCTGTGCTGATGATGCTGTTTGTATCAGGTCGCGGAGTGGCAAGAGAGC 7500
QY 7501 TCGAAATATTTGATGATGCTGACGAGGTTTACCGCAAACTGTTCTGGCGAGCGCCCGT 7560
DB 7501 TCGAAATATTTGATGATGCTGACGAGGTTTACCGCAAACTGTTCTGGCGAGCGCCCGT 7560
QY 7561 TTATCCCGATGGGCGTCTTTGGGCACTTGGCCCGGCTTTGTCATTTGATGCTGCTCA 7620
DB 7561 TTATCCCGATGGGCGTCTTTGGGCACTTGGCCCGGCTTTGTCATTTGATGCTGCTCA 7620
QY 7621 AGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7680
DB 7621 AGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7680
QY 7681 TTTTAAAGAAAATAAGATAGCTTATGCTCCCTATCATCTTCTCATATGCTATGCTGG 7740
DB 7681 TTTTAAAGAAAATAAGATAGCTTATGCTCCCTATCATCTTCTCATATGCTATGCTGG 7740
QY 7741 GATTAGTATTTTGTCTAAATTTGTACTGCTTTGTTTATTCAGATCTCTACGCTATTAC 7800
DB 7741 GATTAGTATTTTGTCTAAATTTGTACTGCTTTGTTTATTCAGATCTCTACGCTATTAC 7800
QY 7801 GAATTCCTTAAATGCTCAACCGGACCCAGCACTTCTGTTCAAGTTATGCTTGGTCTT 7860
DB 7801 GAATTCCTTAAATGCTCAACCGGACCCAGCACTTCTGTTCAAGTTATGCTTGGTCTT 7860
QY 7861 GCCCTAATTTTGTCTCATGTCCTGTGTGTGATGATGATGATGATGATGATGATGAT 7920
DB 7861 GCCCTAATTTTGTCTCATGTCCTGTGTGTGATGATGATGATGATGATGATGATGAT 7920
QY 7921 AAGATTTTGTGTAAGTGAATGTCGAAGTCCCAACATGATGAAAGAAATATATCTAT 7980
DB 7921 AAGATTTTGTGTAAGTGAATGTCGAAGTCCCAACATGATGAAAGAAATATATCTAT 7980
QY 7981 TTTAGTCAATCCCTTGTGTTGTTGATTCATTAATTAATTAATTAATTAATTAATTA 8040
DB 7981 TTTAGTCAATCCCTTGTGTTGTTGATTCATTAATTAATTAATTAATTAATTAATTA 8040
QY 8041 AACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8100
DB 8041 AACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8100
QY 8101 TCTTTTCAATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8160
DB 8101 TCTTTTCAATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8160
QY 8161 GTTAAAGATTTCTTTTACTTAAACACTGCTGCGCATCTTCAATATTTTGAATATCTTGA 8220
DB 8161 GTTAAAGATTTCTTTTACTTAAACACTGCTGCGCATCTTCAATATTTTGAATATCTTGA 8220
QY 8221 CATTAATGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280

```
Db 8221 CATTATGTTCTCTTAAGATTACACAACTTAATTAAGATTTTGTGCTATCAAAATGTT 8280
Qy 8281 TTAGGAAGCTCTTCTTAATAATCTTGAAGAACACAAAGCAGATTTCTTTAAGAGATTA 8340
Db 8281 TTAGGAAGCTCTTCTTAATAATCTTGAAGAACACAAAGCAGATTTCTTTAAGAGATTA 8340
Qy 8341 TGGCTACTAAGGAATCATGAGATTAATGTTAATGGAATAAAGAAAACAATATAT 8400
Db 8341 TGGCTACTAAGGAATCATGAGATTAATGTTAATGGAATAAAGAAAACAATATAT 8400
Qy 8401 TAGGTCCTCAACAAGCAGAGAGATCGATGTTGTAATGGAAGCTAAGAGACTTA 8460
Db 8401 TAGGTCCTCAACAAGCAGAGAGATCGATGTTGTAATGGAAGCTAAGAGACTTA 8460
Qy 8461 CTTTAAAGGTTAATCTGGAGTCTCAAGTCATCCAAACAAATCAATCAAAATATAT 8520
Db 8461 CTTTAAAGGTTAATCTGGAGTCTCAAGTCATCCAAACAAATCAAAATATAT 8520
Qy 8521 TATGTTTGTCTAATGATCTTTTGAAGATGATGATTTGAAGAAATATGAGAGATG 8580
Db 8521 TATGTTTGTCTAATGATCTTTTGAAGATGATGATTTGAAGAAATATGAGAGATG 8580
Qy 8581 AAATTAATTTAGAGATACTTCCGATATCATTTGTCATTTTGTATTAATCTGTA 8640
Db 8581 AAATTAATTTAGAGATACTTCCGATATCATTTGTCATTTTGTATTAATCTGTA 8640
Qy 8641 ATGGTAACACTCAATCATATTAACAAAAGTTTCTCCCATTTTATGTAAGATGACTT 8700
Db 8641 ATGGTAACACTCAATCATATTAACAAAAGTTTCTCCCATTTTATGTAAGATGACTT 8700
Qy 8701 CTTTCTTAATCAATGATTAATCACTTCCATGTAACAGGTCAAACTTAATCACTTTT 8760
Db 8701 CTTTCTTAATCAATGATTAATCACTTCCATGTAACAGGTCAAACTTAATCACTTTT 8760
Qy 8761 GAGAGAGATCATGACGACATTAATTTTGTGCAAGCTGCAAGAGAGAGATCTGTA 8820
Db 8761 GAGAGAGATCATGACGACATTAATTTTGTGCAAGCTGCAAGAGAGATCTGTA 8820
Qy 8821 TTTATGTCAGAGTAGAATGTAATGCGCATTTTAAAGAAAACATATAGAAATATAT 8880
Db 8821 TTTATGTCAGAGTAGAATGTAATGCGCATTTTAAAGAAAACATATAGAAATATAT 8880
Qy 8881 ATCTCTGTTATCTAAACAATCTTCCCATCTTAATAATGCTTACCTTAC 8940
Db 8881 ATCTCTGTTATCTAAACAATCTTCCCATCTTAATAATGCTTACCTTAC 8940
Qy 8941 ACAATTTTACTAGATCTAGTACAAAGTTGAAACAGTTATTTGGACAGAGGAGTAG 9000
Db 8941 ACAATTTTACTAGATCTAGTACAAAGTTGAAACAGTTATTTGGACAGAGGAGTAG 9000
Qy 9001 TATATATGTTGTGAGAACATTAAGGTTATGTTGA CTGATATATGCTTTAAATGTGAAA 9060
Db 9001 TATATATGTTGTGAGAACATTAAGGTTATGTTGA CTGATATATGCTTTAAATGTGAAA 9060
Qy 9061 CAGGTCCTTAATGTTTGTGATGTAAGCAAGTTCTTAACAGTTTCCGAGAGACTAC 9120
Db 9061 CAGGTCCTTAATGTTTGTGATGTAAGCAAGTTCTTAACAGTTTCCGAGAGACTAC 9120
Qy 9121 ACAATTAATGATTAACATATCTGTCAGAAAATGATTTACACATTAAGAAATCTTCT 9180
Db 9121 ACAATTAATGATTAACATATCTGTCAGAAAATGATTTACACATTAAGAAATCTTCT 9180
Qy 9181 TTTTATGCAAGAGACTAGATGAGATGTTCTTCTGATCTATGTTGCTTTTCTTTC 9240
Db 9181 TTTTATGCAAGAGACTAGATGAGATGTTCTTCTGATCTATGTTGCTTTTCTTTC 9240
Qy 9241 TCGTTACATGTTTGTGCTTCAAAAATTAATTAACAGAGCATGTTCCAAATGAT 9300
Db 9241 TCGTTACATGTTTGTGCTTCAAAAATTAATTAACAGAGCATGTTCCAAATGAT 9300
Qy 9301 ATTAAATATTTGAGGTGTTTTCACCAACTTAATATCTTCAAGTCTTAAAAAAC 9360
Db 9301 ATTAAATATTTGAGGTGTTTTCACCAACTTAATATCTTCAAGTCTTAAAAAAC 9360

Db 9301 ATTAAATATTTGAGGTGTTTTCACCAACTTAATATCTTCAAGTCTTAAAAAAC 9360
Qy 9361 GTATATATGTTTAACTCTAACAAAACCTTAATATGTTTCTCTAATACAGGAGGT 9420
Db 9361 GTATATATGTTTAACTCTAACAAAACCTTAATATGTTTCTCTAATACAGGAGGT 9420
Qy 9421 TCTTGAATGGAATTTGGGCTGATTAATCTTTGCGCTGCTGCTCACTTCTTCAAGA 9480
Db 9421 TCTTGAATGGAATTTGGGCTGATTAATCTTTGCGCTGCTGCTCACTTCTTCAAGA 9480
Qy 9481 TGAAGTGAAGAGGTCAATCATCTGTCAAAAGGAACAAGAAAGAAATCTAATAATG 9540
Db 9481 TGAAGTGAAGAGGTCAATCATCTGTCAAAAGGAACAAGAAAGAAATCTAATAATG 9540
Qy 9541 TTTGATGTTGACACACCCCTAGTTGATCATCTGACAGAGCTGTAATCATTTCTAGT 9600
Db 9541 TTTGATGTTGACACACCCCTAGTTGATCATCTGACAGAGCTGTAATCATTTCTAGT 9600
Qy 9601 ATCCCATTTAATTAATTTCAATTAACATATGTAATGTTCTGTTGAGCTGCAAGT 9660
Db 9601 ATCCCATTTAATTAATTTCAATTAACATATGTAATGTTCTGTTGAGCTGCAAGT 9660
Qy 9661 CATGACTTAATTTTGAATGATTTGGCTCATTTGCTGCAATCAATTAATAAT 9720
Db 9661 CATGACTTAATTTTGAATGATTTGGCTCATTTGCTGCAATCAATTAATAAT 9720
Qy 9721 GGTGTTGTACAAATGATGATGAGAGCGAGGTGTTTGAACACCTTTTCAACAAAAT 9780
Db 9721 GGTGTTGTACAAATGATGATGAGAGCGAGGTGTTTGAACACCTTTTCAACAAAAT 9780
Qy 9781 CTATATCTTTCAACAATGAAACCTTGAATCCCTTGAAGAGCAACATCTCTT 9840
Db 9781 CTATATCTTTCAACAATGAAACCTTGAATCCCTTGAAGAGCAACATCTCTT 9840
Qy 9841 GAATATGCTATGTTTCCATGATCTGATGAAACATGATGAATGAGTAATATC 9900
Db 9841 GAATATGCTATGTTTCCATGATCTGATGAAACATGATGAATGAGTAATATC 9900
Qy 9901 CATGCAAGTTTATTAATGTTAATTTCAATATGAGAACTTTGATTAATCTTACAC 9960
Db 9901 CATGCAAGTTTATTAATGTTAATTTCAATATGAGAACTTTGATTAATCTTACAC 9960
Qy 9961 ACATTCCTGAAGTAATGTCAGTTTGTGCTTGAAGAGGACCTAATTTTCCATGGG 10020
Db 9961 ACATTCCTGAAGTAATGTCAGTTTGTGCTTGAAGAGGACCTAATTTTCCATGGG 10020
Qy 10021 GGGGGGGGTGAATAGCGGTTTATTAACAATGATATTTGAGAAATCTTAATGTGGA 10080
Db 10021 GGGGGGGGTGAATAGCGGTTTATTAACAATGATATTTGAGAAATCTTAATGTGGA 10080
Qy 10081 ATTAACTAGGTGAATATTTTTCCAATTAAGAGGTGCTTTATGACTCACATGTACA 10140
Db 10081 ATTAACTAGGTGAATATTTTTCCAATTAAGAGGTGCTTTATGACTCACATGTACA 10140
Qy 10141 TCAAGGATCAATCATATGATGATGACATTCGACATTAATCAAGGTGATCGG 10200
Db 10141 TCAAGGATCAATCATATGATGATGACATTCGACATTAATCAAGGTGATCGG 10200
Qy 10201 CCAACACACACACGACACACATTCACACACAAATCAAGTGAAGAGGAA 10260
Db 10201 CCAACACACACACGACACACATTCACACACAAATCAAGTGAAGAGGAA 10260
Qy 10261 GTCATACAAATCAAAAATGCTGCTAGCGGAGAGAAATGAAAACATGAAAGAAATGA 10320
Db 10261 GTCATACAAATCAAAAATGCTGCTAGCGGAGAGAAATGAAAACATGAAAGAAATGA 10320
Qy 10321 AAAACGTGACGACACATCTGACACGACAAACATCTGTGACACACAAA 10380
Db 10321 AAAACGTGACGACACATCTGACACGACAAACATCTGTGACACACAAA 10380
Qy 10381 CTGCGAGAAAAGTTCTAATAAATGCGGCTTGAAGAGAAACGAGCTGCAAGATGTC 10440
Db 10381 CTGCGAGAAAAGTTCTAATAAATGCGGCTTGAAGAGAAACGAGCTGCAAGATGTC 10440
```

Qy	10441	CATCATCGATCCAAACCACTAAGTCAATATCTGGTTTTCATCTCGAAGATCAATCCG	10500
Db	10441	CATCATCGATCCAAACCACTAAGTCAATATCTGGTTTTCATCTCGAAGATCAATCCG	10500
Qy	10501	AGCAAACTCCGAGTAAATGCTTTTATTAAGGTAACGATTCAAAAAATGCCACATCATGAG	10560
Db	10501	AGCAAACTCCGAGTAAATGCTTTTATTAAGGTAACGATTCAAAAAATGCCACATCATGAG	10560
Qy	10561	TTATGACCAATTTAGACCAAGACCTAGATTTTATCCAAAGCTGAGACGGGTACTCTAGA	10620
Db	10561	TTATGACCAATTTAGACCAAGACCTAGATTTTATCCAAAGCTGAGACGGGTACTCTAGA	10620
Qy	10621	AGTACATCCAAATTGAAGTCAATCCCACTTGGCTTCATCAATTAATAGTTGACATAGTACG	10680
Db	10621	AGTACATCCAAATTGAAGTCAATCCCACTTGGCTTCATCAATTAATAGTTGACATAGTACG	10680
Qy	10681	GTCATATGGCGAGTAAATGACATAGACGGCAGATGTAGTTAGCTGACGTGACAGA	10740
Db	10681	GTCATATGGCGAGTAAATGACATAGACGGCAGATGTAGTTAGCTGACGTGACAGA	10740
Qy	10741	GCCGTGCGCCACCACTGACGAGATGTTTATGAGGAGAGAAAGTATGGCTCCACCAAC	10800
Db	10741	GCCGTGCGCCACCACTGACGAGATGTTTATGAGGAGAGAAAGTATGGCTCCACCAAC	10800
Qy	10801	ATCCCAAGTTTGAACATTTCTAGAGCCCTTACCATCTCAAAAGCAGCAATTGATGAC	10860
Db	10801	ATCCCAAGTTTGAACATTTCTAGAGCCCTTACCATCTCAAAAGCAGCAATTGATGAC	10860
Qy	10861	TATCTGATATGACGACAAATCCATGTCCTGCTCATCTGCTCATTTGCTATGACATAC	10920
Db	10861	TATCTGATATGACGACAAATCCATGTCCTGCTCATCTGCTCATTTGCTATGACATAC	10920
Qy	10921	TACCTGCGAAAGCGGAGTTCAGAGCCCGACAGAGCTGGCGCGCCG	10966
Db	10921	TACCTGCGAAAGCGGAGTTCAGAGCCCGACAGAGCTGGCGCGCCG	10966

RESULT 3
AB024006
LOCUS
DEFINITION
10966 bp DNA linear PLN 27-NOV-1999
Hordeum vulgare naat-B and naat-A genes for nicotianamine
amino transferase, complete cds.

AB024006
AB024006.1 GI:6469088
nicotianamine amino transferase.
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE
AUTHORS
1 (sites)
Takahashi, M., Yamaguchi, H., Nakanishi, H., Shiota, T.,
Nishizawa, N. K. and Mori, S.

TITLE
Cloning two genes for nicotianamine amino transferase, a critical
enzyme in iron acquisition (Strategy II) in graminaceous plants
JOURNAL
Plant Physiol. 121 (3), 947-956 (1999)

PUBMED
20027437
10557244
2 (bases 1 to 10966)
Mori, S., Takahashi, M. and Nakanishi, H.

REFERENCE
AUTHORS
Direct Submission
Submitted (22-FEB-1999) Satoshi Mori, University of Tokyo, Dept. of
Applied Biological Chemistry, Yayoi 1-1, Bunkyo-ku, Tokyo 113-8657,
Japan (E-mail: asamori@hongo.ecc.u-tokyo.ac.jp,
Tel: 81-3-3812-2111 (ex. 5106), Fax: 81-3-3812-0544)

FEATURES
source
1..10966
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="genomic DNA"
/sub_species="Igr1"
/db_xref="taxon:112509"
/clone_1nb="lambda FIX11"

gene	join(650..1299,1483..1825,1922..2140,2244..2303, 2761..2916,3263..3356,3735..3868)
CDS	/gene="naat-B" join(650..1299,1483..1825,1922..2140,2244..2303, 2761..2916,3263..3356,3735..3868) /gene="naat-B" /codon_start=1 /product="nicotianamine amino transferase" /protein_id="BA087054.1" /db_xref="GI:6469089" /translation="MATTROSDGVNANGLAVAAAANGKSNHGVAAVANGKNGHGD ADANGSNHGVANANGKSNHAEATANGHGAATANGKTNHRESNGHAEADANSE SNEHADSDANGESNGHAAAEAEAEVMEFAGADGVLAATGAMNSIAIRYKISA SVOEKSPRPVLPLAGDPSVFPARTAVEEDVAAALPTGCTNCTPAGVPLAASA VAEHSQVYMLSDADVFLTAGTAEVILVLAOTGANTLLPBGVYENBARA FNRLEVRHFDLIPDKMELIDISLESIADKNTAMVITPNPNBGSVSYDHLSTVAE VAKRLGILVADSVYKLVGSPFLPMKVFHTTTLVLSISLSKSNIVGMLGVYA VDPKRLIDETKISTITNYLVNSTDPAFTQLAALQIILESTEDCKKLAKESVILCPG EICVKIKENKXITCPHKEGSMFVWVKLNLLEBIDIDIDICCKKLAKESVILCPG SVLGMAWVRITPACVPSLQDGLGRKISFCQPKRNSDDC" join(6518..6897,7029..7371,7479..7697,7784..7843, 8285..8440,8738..8831,9414..9547) /gene="naat-A" join(6518..6897,7029..7371,7479..7697,7784..7843, 8285..8440,8738..8831,9414..9547) /gene="naat-A" /codon_start=1 /product="nicotianamine amino transferase" /protein_id="BA087055.1" /db_xref="GI:6469090" /translation="MVHQSNGHEAAAAAANGKSNHAAAAANGKSNHAAAAAENKF ARKKDIILATITGAKNSIRAIRKISASVESGPRPLPLAGDPSVFPARTAVEED AVAAALPTGCTNCTPAGVPLAASAVALHLSQGVYKLSADVPFLAGTQALVYI PVLAOTGANTLLPBGVYENBARAFAENFLVLPDKMEIDISLESIADKNT TMVITPNPNBGSVSYDHLAKVAEVARLGLIVLADSVYKLVGSPFLPMKVFHTT HAPVLSISLSKSNIVPGRMLGVYAVYPTKILKTKISTITNYLVNSTDPAFTYQ EALPKLENTKADFFRRLITGLNESSLEITREIKENKXITCPHKEGSMFVWVKLNL LLEIHIDIDINFCCKKLAKESVILCPGSLVGMENWVRITPACVPSLQDGLERKVSFCQ RNKKNSINGC"

ORIGIN

Query Match	100.0%; Score 10966; DB 8; Length 10966;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 10966; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGATCCCATTCGATGATGATGATGATGATCAACGAAAGAAAGATGCGATGCGC 60
Db	1 CTGATCCCATTCGATGATGATGATGATGATCAACGAAAGAAAGATGCGATGCGC 60
Qy	61 CTGTGTGTCATCCCTCACTGGCTTGGCGAATGCGCATACCGAGTTAGTGAAGTGT 120
Db	61 CTGTGTGTCATCCCTCACTGGCTTGGCGAATGCGCATACCGAGTTAGTGAAGTGT 120
Qy	121 TTAGCATGATGTCGTGCGGCACTGCCAAGAAAATCGCGTGCAGCGGACTGCGAGAGTT 180
Db	121 TTAGCATGATGTCGTGCGGCACTGCCAAGAAAATCGCGTGCAGCGGACTGCGAGAGTT 180
Qy	181 GAGCGATGATGCTTGTGTATGAGCGAGCTGAGTGGTGTCACTAATGAAACCAATCA 240
Db	181 GAGCGATGATGCTTGTGTATGAGCGAGCTGAGTGGTGTCACTAATGAAACCAATCA 240
Qy	241 GCATTGGGTGATGCGAGTGAAGATCATGCTTCTGCGTCCGATCCGCTTAATCTTT 300
Db	241 GCATTGGGTGATGCGAGTGAAGATCATGCTTCTGCGTCCGATCCGCTTAATCTTT 300
Qy	301 TTCTCCCAATTAATAAGGGATAGATGATGATGATGATGATGATGATGATGATGATG 360
Db	301 TTCTCCCAATTAATAAGGGATAGATGATGATGATGATGATGATGATGATGATGATG 360
Qy	361 ATGAAACCAAGGAGCGCCGAAATATGCGGGGATATGATGATGATGATGATGATGATG 420
Db	361 ATGAAACCAAGGAGCGCCGAAATATGCGGGGATATGATGATGATGATGATGATGATG 420

Qy 421 CACGCCCTGACGTGGACGAGCGGCTGCGACGAGCGGTCAGACGCGGCTGCTGCT 480
Db 421 CACGCCCTGACGTGGACGAGCGGCTGCGACGAGCGGTCAGACGCGGCTGCTGCT 480
Qy 481 TCTTGTGCTCCGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 TCTTGTGCTCCGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 541 ACCGCTACTCAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 ACCGCTACTCAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 601 GTAGTCTAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 GTAGTCTAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 AGCGCAGAGCGACGAGTGGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCT 720
Db 661 AGCGCAGAGCGACGAGTGGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCT 720
Qy 721 GAGCAACGCGACGAGTGGCGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCT 780
Db 721 GAGCAACGCGACGAGTGGCGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 TGCGGACGCGACGAGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 TGCGGACGCGACGAGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 841 CAACGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 841 CAACGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy 901 CAACGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 901 CAACGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 961 CGAGCATGCGAGGACTCGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 CGAGCATGCGAGGACTCGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 AGAGGAGGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 AGAGGAGGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 GACGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 1081 GACGCGCGCGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy 1141 GAAAGGAGGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 GAAAGGAGGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1201 CGGACGCGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 CGGACGCGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy 1261 CTGCTGACCGCGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 1261 CTGCTGACCGCGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Qy 1321 GAAAGGAGGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db 1321 GAAAGGAGGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Qy 1381 TTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 1381 TTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Qy 1441 GACCTTCTCCGACGCTTATACCAAAACAAACAAACAAACAAACAAACAAACAAAC 1500
Db 1441 GACCTTCTCCGACGCTTATACCAAAACAAACAAACAAACAAACAAACAAACAAAC 1500
Qy 1501 GTCCAGAGGCGGTGCGGTATGCTATCGAGCGACGAGCTTCTGCTGCTGCTGCTGCT 1560

Db 1501 GTCCAGAGGCGGTGCGGTATGCTATCGAGCGACGAGCTTCTGCTGCTGCTGCTGCT 1560
Qy 1561 CCAGGCGATGAGGCTATATCCCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db 1561 CCAGGCGATGAGGCTATATCCCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Qy 1621 CCCAGAGGCGAGCTATCCCAACCTACAGAGCGCGCGCGCTTCAACAGCTGAGAGTCCG 1680
Db 1621 CCCAGAGGCGAGCTATCCCAACCTACAGAGCGCGCGCGCTTCAACAGCTGAGAGTCCG 1680
Qy 1681 GCATTTGACCTCATCCCGGACAAAGGAGGAGAGTGAATGACATGCACTGCTGAATTCAT 1740
Db 1681 GCATTTGACCTCATCCCGGACAAAGGAGGAGAGTGAATGACATGCACTGCTGAATTCAT 1740
Qy 1741 CGCGGACAAAGAACACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1741 CGCGGACAAAGAACACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Qy 1801 TTACTCTGACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 TTACTCTGACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy 1861 GTTCAGTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 GTTCAGTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 1921 GGTGCGGAGGCTGCGGAAAGGCTGCGGAAATTTGGTGAATGCTGACGAGGTAACGCGGAA 1980
Db 1921 GGTGCGGAGGCTGCGGAAAGGCTGCGGAAATTTGGTGAATGCTGACGAGGTAACGCGGAA 1980
Qy 1981 GGTGCGGAGGCTGCGGAAAGGCTGCGGAAATTTGGTGAATGCTGACGAGGTAACGCGGAA 2040
Db 1981 GGTGCGGAGGCTGCGGAAAGGCTGCGGAAATTTGGTGAATGCTGACGAGGTAACGCGGAA 2040
Qy 2041 GCTGCTCATAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2041 GCTGCTCATAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Qy 2101 AGCGGTGACGACCCCGGAAAGATCTTACAGGAACTTAAGGTAATCTTAATATCA 2160
Db 2101 AGCGGTGACGACCCCGGAAAGATCTTACAGGAACTTAAGGTAATCTTAATATCA 2160
Qy 2161 TTCTTTCAATGCTATCTAAGGTAATTAATTAAGTACTGATCAATATATATGCTAAT 2220
Db 2161 TTCTTTCAATGCTATCTAAGGTAATTAATTAAGTACTGATCAATATATATGCTAAT 2220
Qy 2221 TTGTAAGTCAATTTTGTGTAATCTTACATCAATTAACAAATTAACCTCAATGCTGCA 2280
Db 2221 TTGTAAGTCAATTTTGTGTAATCTTACATCAATTAACAAATTAACCTCAATGCTGCA 2280
Qy 2281 CAGACCGACGACCTTCAATCAAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2281 CAGACCGACGACCTTCAATCAAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Qy 2341 GTCTTGTGTAATTAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 2341 GTCTTGTGTAATTAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Qy 2401 CACGTGATGTCAGAGTCAAGTCTTCAATTAAGTGAATTAAGTCAAGTCAAGTCAAT 2460
Db 2401 CACGTGATGTCAGAGTCAAGTCTTCAATTAAGTGAATTAAGTCAAGTCAAGTCAAT 2460
Qy 2461 ACATTTGATGTAATTTTACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Db 2461 ACATTTGATGTAATTTTACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Qy 2521 CTGAACCTACTGTAATTAATACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Db 2521 CTGAACCTACTGTAATTAATACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Qy 2581 CATGCAATATCTTTGGGCTCAAGTATGCAATATGGAACCTCTTTTGAATAATGAT 2640

Db 2581 CAGCCATATTCCTTTGGGTCAGATGACGATAGTTGGAACTCTTTTATGAAAAATAGAT 2640
Qy 2641 ACATGTACTAGATGATACCATTTTATTAAGAAATTCATATTTTGAATCCTGATGGT 2700
Db 2641 ACATGTACTAGATGATACCATTTTATTAAGAAATTCATATTTTGAATCCTGATGGT 2700
Qy 2701 ATGTCTCTGTGATTCACACGATTTAATTGGTTTATTTTGAATCCTGATGGT 2760
Db 2701 ATGTCTCTGTGATTCACACGATTTAATTGGTTTATTTTGAATCCTGATGGT 2760
Qy 2761 GCACCTCTCTGATTCGATTCGAGAAACAAAGGAATTTCTTAAGGCGATATAGGT 2820
Db 2761 GCACCTCTCTGATTCGATTCGAGAAACAAAGGAATTTCTTAAGGCGATATAGGT 2820
Qy 2821 CTGCTAAAGAAATCATCGATATGCTACAAACAAATTAAGAAAAATATACATTTA 2880
Db 2821 CTGCTAAAGAAATCATCGATATGCTACAAACAAATTAAGAAAAATATACATTTA 2880
Qy 2881 TGTCTCTCAAGGACGAAAGATCAATGTTTGTCTAGGCTATTTTGTGAATGAAA 2940
Db 2881 TGTCTCTCAAGGACGAAAGATCAATGTTTGTCTAGGCTATTTTGTGAATGAAA 2940
Qy 2941 AAATCTTAGGGAGTGTGATCATATACTATTTATATAGATTAATCGGGACGAA 3000
Db 2941 AAATCTTAGGGAGTGTGATCATATACTATTTATATAGATTAATCGGGACGAA 3000
Qy 3001 ATGCATCAACATATTAATCTCAAAATTCAAATTCATCTCCGATACATATTT 3060
Db 3001 ATGCATCAACATATTAATCTCAAAATTCAAATTCATCTCCGATACATATTT 3060
Qy 3061 TGAAGATGATGATTTTAAAGATTAAGACGAGCTTAAGTTTGTCTACATATCAT 3120
Db 3061 TGAAGATGATGATTTTAAAGATTAAGACGAGCTTAAGTTTGTCTACATATCAT 3120
Qy 3121 CTGATATTCCTTTGATTCATCTTTTGTATATCTGGAATGTATGCTCAATCATAT 3180
Db 3121 CTGATATTCCTTTGATTCATCTTTTGTATATCTGGAATGTATGCTCAATCATAT 3180
Qy 3181 ACACAAAATTCATGATGATTTCTAGAAAGATTTATTTCTGTAACCATGACCTCC 3240
Db 3181 ACACAAAATTCATGATGATTTCTAGAAAGATTTATTTCTGTAACCATGACCTCC 3240
Qy 3241 GTATTAATTCATGATGATTTCTAGAAAGATTTATTTCTGTAACCATGACCTCC 3300
Db 3241 GTATTAATTCATGATGATTTCTAGAAAGATTTATTTCTGTAACCATGACCTCC 3300
Qy 3301 TGACATGATTTTGTGCAAGCTCGCAAAAGAAAGATTCATATTCGAGATGACCTA 3360
Db 3301 TGACATGATTTTGTGCAAGCTCGCAAAAGAAAGATTCATATTCGAGATGACCTA 3360
Qy 3361 GAATCCATGTTGATTTTGTGCAAGCTCGCAAAAGAAAGATTCATATTCGAGATGAC 3420
Db 3361 GAATCCATGTTGATTTTGTGCAAGCTCGCAAAAGAAAGATTCATATTCGAGATGAC 3420
Qy 3421 CATTAATGATTTACATATTTATGTAAGAAATTTTATTAAGAAATTTCTGTCG 3480
Db 3421 CATTAATGATTTACATATTTATGTAAGAAATTTTATTAAGAAATTTCTGTCG 3480
Qy 3481 AAGACTTTTAAACAAATTTCTGTAAGAAATTTATTAAGAAATTTCTGTCG 3540
Db 3481 AAGACTTTTAAACAAATTTCTGTAAGAAATTTATTAAGAAATTTCTGTCG 3540
Qy 3541 TTGACCTGATGTCCTGTCATGTTGTGCTCTCAAAAATAATGACGACATGTT 3600
Db 3541 TTGACCTGATGTCCTGTCATGTTGTGCTCTCAAAAATAATGACGACATGTT 3600
Qy 3601 TCAAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 3660
Db 3601 TCAAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 3660
Qy 3661 TTTTGTAGTCAACTTATATTTTGTGACATTTCTCAAAAACAAACTTATATATGTG 3720
Db 3661 TTTTGTAGTCAACTTATATTTTGTGACATTTCTCAAAAACAAACTTATATATGTG 3720

Qy 3721 CCTCAAAATGATGAGGATGTTCTTGAATGCAAACTGGGTCGACATTACTTTGCTTG 3780
Db 3721 CCTCAAAATGATGAGGATGTTCTTGAATGCAAACTGGGTCGACATTACTTTGCTTG 3780
Qy 3781 TGTTCATCTCTTCAAGATGCTCGGAAGGATCAATCTCTGTCAAAGAA 3840
Db 3781 TGTTCATCTCTTCAAGATGCTCGGAAGGATCAATCTCTGTCAAAGAA 3840
Qy 3841 GAAAGAAATTCGAGGATGATGCTAGTTATGATGATGATGATGATGATGATGATG 3900
Db 3841 GAAAGAAATTCGAGGATGATGCTAGTTATGATGATGATGATGATGATGATGATG 3900
Qy 3901 CAGATCCCATGATATCTTTCAATTAATGAACTTTTATGTTCTGATGATGATGATG 3960
Db 3901 CAGATCCCATGATATCTTTCAATTAATGAACTTTTATGTTCTGATGATGATGATG 3960
Qy 3961 CAACATCTCTTGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Db 3961 CAACATCTCTTGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Qy 4021 GTTAGTAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
Db 4021 GTTAGTAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
Qy 4081 GGGGGGGGGGCTTGTGATTAATCTTTAAGTACAGTTCTCAAGTTATGTCAAAGCA 4140
Db 4081 GGGGGGGGGGCTTGTGATTAATCTTTAAGTACAGTTCTCAAGTTATGTCAAAGCA 4140
Qy 4141 CTTTGTAAATTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4200
Db 4141 CTTTGTAAATTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4200
Qy 4201 AAGCAAAACCTTACCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
Db 4201 AAGCAAAACCTTACCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
Qy 4261 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4320
Db 4261 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4320
Qy 4321 CTACATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4380
Db 4321 CTACATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4380
Qy 4381 CTCTGTGCAAAATTCGGAACCAAGATTAATCATCTCAAGCTGGAAGCAACCAAAATTTG 4440
Db 4381 CTCTGTGCAAAATTCGGAACCAAGATTAATCATCTCAAGCTGGAAGCAACCAAAATTTG 4440
Qy 4441 AAGAGGAAACCTTCAAAAGACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4500
Db 4441 AAGAGGAAACCTTCAAAAGACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4500
Qy 4501 TCGCAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4560
Db 4501 TCGCAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4560
Qy 4561 CGATTAAGTAAATTCGGAACCAAGATTAATCATCTCAAGCTGGAAGCAACCAAAATTTG 4620
Db 4561 CGATTAAGTAAATTCGGAACCAAGATTAATCATCTCAAGCTGGAAGCAACCAAAATTTG 4620
Qy 4621 AATTTGAAATTTGTAACCAATTTTAAACGGGATTCCTGGAACATTTTCAAAATTTG 4680
Db 4621 AATTTGAAATTTGTAACCAATTTTAAACGGGATTCCTGGAACATTTTCAAAATTTG 4680
Qy 4681 GATCAAAATTTTAAACGATTTTCAAAATTTTAAACGGGATTCCTGGAACATTTTCAAAATTTG 4740
Db 4681 GATCAAAATTTTAAACGATTTTCAAAATTTTAAACGGGATTCCTGGAACATTTTCAAAATTTG 4740
Qy 4741 GTTCAACAAATTTTGAACCTTTTAAATTTAGCAGAAACATTTTGAATTTCTAATATTTT 4800
Db 4741 GTTCAACAAATTTTGAACCTTTTAAATTTAGCAGAAACATTTTGAATTTCTAATATTTT 4800

4801 CGAATTTGGAACTTTTTCTATTCTGACAAAAATTGAAATATGCAAGCGTAATTTGGA 4860
4801 CGAATTTGGAACTTTTTCTATTCTGACAAAAATTGAAATATGCAAGCGTAATTTGGA 4860
4861 ATAAATTTGGAAAAATGCGATTTTTTGAATTTCTGAAATATTTGAAAAACAAAAAA 4920
4861 ATAAATTTGGAAAAATGCGATTTTTTGAATTTCTGAAATATTTTGAAAAAACAAAAAA 4920
4921 CTTTAAAAAGTAAAAATAAAAATAAAAATAGAAATATGAAATATGCAAAAAATA 4980
4921 CTTTAAAAAGTAAAAATAAAAATAAAAATAGAAATATGAAATATGCAAAAAATA 4980
4981 AAGAAATCCGAGAAAGCCAACTGGGAATGACATGGAATAACCCAGCGCTCGCGCCG 5040
4981 AAGAAATCCGAGAAAGCCAACTGGGAATGACATGGAATAACCCAGCGCTCGCGCCG 5040
5041 ACTGTGTAAGCTATTAAGTAGAGCCGCGCCAAAGCTCGTCTCATCATACCTGTGCGA 5100
5041 ACTGTGTAAGCTATTAAGTAGAGCCGCGCCAAAGCTCGTCTCATCATACCTGTGCGA 5100
5101 AACCCCGCAATTTGCTGCACTATGCGCGCAATAGGCTTTTCAGAGCTCTGTCTCC 5160
5101 AACCCCGCAATTTGCTGCACTATGCGCGCAATAGGCTTTTCAGAGCTCTGTCTCC 5160
5161 GGTATGAGGTCAATTTGACACCCCTCTCACTTGGGCGAGGCTATTAATCTTTTTC 5220
5161 GGTATGAGGTCAATTTGACACCCCTCTCACTTGGGCGAGGCTATTAATCTTTTTC 5220
5221 TTCTTTGCACTCAAGTTACTACGCGCACTTAATTTTGAAGCGACCAACCGGTTTTGT 5280
5221 TTCTTTGCACTCAAGTTACTACGCGCACTTAATTTTGAAGCGACCAACCGGTTTTGT 5280
5281 GAAGGTTCTAGAAATCTAACCATTTTGGGAAGCTTCTAGAAAGCTATGATGTTTTCTT 5340
5281 GAAGGTTCTAGAAATCTAACCATTTTGGGAAGCTTCTAGAAAGCTATGATGTTTTCTT 5340
5341 TGACATGTAATTAATTTGTTGTTTTCTTTCAATGACAACTTTTCAATTCAT 5400
5341 TGACATGTAATTAATTTGTTGTTTTCTTTCAATGACAACTTTTCAATTCAT 5400
5401 GATTTTGTGAAACCTGTGATTTTGAATCGGTATTTTTCCTAAATCCGTTT 5460
5401 GATTTTGTGAAACCTGTGATTTTGAATCGGTATTTTTCCTAAATCCGTTT 5460
5461 GAAATTAACCTGTGATTTTCCGAATTAATGAACTTTGTTGCAAGATGATGATCT 5520
5461 GAAATTAACCTGTGATTTTCCGAATTAATGAACTTTGTTGCAAGATGATGATCT 5520
5521 TTTCAATGAGCGAATTTTCTAAATATGCACTATTTTCAATTCATTAAGCTTCC 5580
5521 TTTCAATGAGCGAATTTTCTAAATATGCACTATTTTCAATTCATTAAGCTTCC 5580
5581 TTTTAACTGTAATCTATGCACTTTGTAATCTTTTAAATTTTCTTTTAAATGA 5640
5581 TTTTAACTGTAATCTATGCACTTTGTAATCTTTTAAATTTTCTTTTAAATGA 5640
5641 TTTTAACTGTAATCTATGCACTTTGTAATCTTTTAAATTTTCTTTTAAATGA 5700
5641 TTTTAACTGTAATCTATGCACTTTGTAATCTTTTAAATTTTCTTTTAAATGA 5700
5701 AACCATGACTCGCTGAAACAGGCAAGGCTTCAATGATGAGGTGCTTAATACAGCG 5760
5701 AACCATGACTCGCTGAAACAGGCAAGGCTTCAATGATGAGGTGCTTAATACAGCG 5760
5761 AACCATGACTCGCTGAAACAGGCAAGGCTTCAATGATGAGGTGCTTAATACAGCG 5820
5761 AACCATGACTCGCTGAAACAGGCAAGGCTTCAATGATGAGGTGCTTAATACAGCG 5820
5821 AATATTCGATTAATTAAGGGAATTAATTAATTAATTAATTAATTAATTAATTAAT 5880
5821 AATATTCGATTAATTAAGGGAATTAATTAATTAATTAATTAATTAATTAATTAAT 5880
5881 AATATTCGATTAATTAAGGGAATTAATTAATTAATTAATTAATTAATTAATTAAT 5940
5881 AATATTCGATTAATTAAGGGAATTAATTAATTAATTAATTAATTAATTAATTAAT 5940

5881 AATTTGTTGTTGGGAGAACTAATCTGAAGAATCTAATCTAGTTCTCCGCAATTAACAATA 5940
5941 TTATGATTTGGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6000
5941 TTATGATTTGGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6000
6001 ACCCAAGAAAGTCTAGAGTTTGTGTTTGAAGACCGGAAGATGATGATGATGATGATG 6060
6001 ACCCAAGAAAGTCTAGAGTTTGTGTTTGAAGACCGGAAGATGATGATGATGATGATG 6060
6061 ATTAGTCCATCTTTTCCCAAGATTAAGAAAGATGAATTTGATTAATTAATTAATTAAT 6120
6061 ATTAGTCCATCTTTTCCCAAGATTAAGAAAGATGAATTTGATTAATTAATTAATTAAT 6120
6121 GAGATATCAATTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6180
6121 GAGATATCAATTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6180
6181 AGCGGAATTTGGAATTTGAGTGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 6240
6181 AGCGGAATTTGGAATTTGAGTGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 6240
6241 TGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6300
6241 TGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6300
6301 GAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6360
6301 GAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6360
6361 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6420
6361 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6420
6421 GCGGTTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 6480
6421 GCGGTTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 6480
6481 CCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 6540
6481 CCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 6540
6541 CCGGGAAGCG 6600
6541 CCGGGAAGCG 6600
6601 CCGGGAAGCG 6660
6601 CCGGGAAGCG 6660
6661 GAGCGGATCTCTGCG 6720
6661 GAGCGGATCTCTGCG 6720
6721 CAGCGGATCTCTGCG 6780
6721 CAGCGGATCTCTGCG 6780
6781 GTCGTTCTCTGCG 6840
6781 GTCGTTCTCTGCG 6840
6841 GCGGATCTCTGCG 6900
6841 GCGGATCTCTGCG 6900
6901 ACATTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAG 6960
6901 ACATTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAGTTTACAG 6960
6961 CCGCGGTTCTCTGCG 7020
6961 CCGCGGTTCTCTGCG 7020

Dh 6961 CCGGCGGTGTTCTTCCCGGTGCGCTTCAAAATTTTAACTTCTATTAAGTACTTATATAA 7020
Qy 7021 ACAAAGCGCGCGGTAGAGACCTTGTCAAGAGGCGGCTTCAAGACTATGCGCGGAC 7080
Dh 7021 ACAAAGCGCGCGGTAGAGACCTTGTCAAGAGGCGGCTTCAAGACTATGCGCGGAC 7080
Qy 7081 GACGCTTCTCTACCGCGCGGGAACCTCAGGCGATCGAAGTCAATATCCCGGCTGAGCC 7140
Dh 7081 GACGCTTCTCTACCGCGCGGGAACCTCAGGCGATCGAAGTCAATATCCCGGCTGAGCC 7140
Qy 7141 CAGAGCTCGCGCGCGCAATATCTGCTTCCCGCGCAAGCTATCCAAATTAAGAGCGCGA 7200
Dh 7141 CAGAGCTCGCGCGCGCAATATCTGCTTCCCGCGCAAGCTATCCAAATTAAGAGCGCGA 7200
Qy 7201 GCGGCAATTCACAAAGCTGGAAGTCCGGCACTTCGACCTCATCCCGACAAAGGGTGGAG 7260
Dh 7201 GCGGCAATTCACAAAGCTGGAAGTCCGGCACTTCGACCTCATCCCGACAAAGGGTGGAG 7260
Qy 7261 ATCGACATCGACTCGCTGGAATCCATCGCCGACAAAGAACACACCGGATGCTCATTA 7320
Dh 7261 ATCGACATCGACTCGCTGGAATCCATCGCCGACAAAGAACACACCGGATGCTCATTA 7320
Qy 7321 AACCCAAACATCCGTCGCGGCAAGGCTTACTCTTACGACCATCTGGCCAGGTTTGAT 7380
Dh 7321 AACCCAAACATCCGTCGCGGCAAGGCTTACTCTTACGACCATCTGGCCAGGTTTGAT 7380
Qy 7381 CCATGACATCTCTGCTGCTGTTGATCGACCGGCTGTTGAACATAGTATGATGCGT 7440
Dh 7381 CCATGACATCTCTGCTGCTGTTGATCGACCGGCTGTTGAACATAGTATGATGCGT 7440
Qy 7441 TTGCTAATCGTGTGCTGATGATGCTGTTGTTATCAAGTGGCGAGGTGGCAAGAAAGC 7500
Dh 7441 TTGCTAATCGTGTGCTGATGATGCTGTTGTTATCAAGTGGCGAGGTGGCAAGAAAGC 7500
Qy 7501 TCGGAATATGATGATGATGCTGCTACAGAGTTTACGGCAAACTGTTCTGGGAGCGCCCGT 7560
Dh 7501 TCGGAATATGATGATGATGCTGCTACAGAGTTTACGGCAAACTGTTCTGGGAGCGCCCGT 7560
Qy 7561 TTATCCCGATGGGCGTCTTTGGGCACTTGGCCGCGTCTTGTCCATTGATGATCTGTCCA 7620
Dh 7561 TTATCCCGATGGGCGTCTTTGGGCACTTGGCCGCGTCTTGTCCATTGATGATCTGTCCA 7620
Qy 7621 AGTCGTGATGATGCTGCTGATGCGGACTTGGATGGGCGGTGACGACCCCAAAAGA 7680
Dh 7621 AGTCGTGATGATGCTGCTGATGCGGACTTGGATGGGCGGTGACGACCCCAAAAGA 7680
Qy 7681 TTTTAGAGAAAATGATGATGCTTAACTGCTCCCTATCATTTCTTCATATGCTATGAGG 7740
Dh 7681 TTTTAGAGAAAATGATGATGCTTAACTGCTCCCTATCATTTCTTCATATGCTATGAGG 7740
Qy 7741 GATTAGTATTTTGTGCTAAATTTGATGCTGTTTATTCAGATCTCTACGCTATTTAC 7800
Dh 7741 GATTAGTATTTTGTGCTAAATTTGATGCTGTTTATTCAGATCTCTACGCTATTTAC 7800
Qy 7801 GAAATTAATTAATGCTCAAGGACCCAGCAACCTTGTCAAGTTAGTCTTTGGTCTT 7860
Dh 7801 GAAATTAATTAATGCTCAAGGACCCAGCAACCTTGTCAAGTTAGTCTTTGGTCTT 7860
Qy 7801 GAAATTAATTAATGCTCAAGGACCCAGCAACCTTGTCAAGTTAGTCTTTGGTCTT 7860
Dh 7801 GAAATTAATTAATGCTCAAGGACCCAGCAACCTTGTCAAGTTAGTCTTTGGTCTT 7860
Qy 7861 GCGCATATTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7920
Dh 7861 GCGCATATTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7920
Qy 7921 AGAGTTTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 7980
Dh 7921 AGAGTTTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 7980
Qy 7981 TTTAGTCAATCCCTTTGTTGTTGATTCATTAATAATAATAATAATAATAATAATAATAATA 8040
Dh 7981 TTTAGTCAATCCCTTTGTTGTTGATTCATTAATAATAATAATAATAATAATAATAATAATA 8040
Qy 8041 AACATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 8100
Dh 8041 AACATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 8100

Qy 8101 TCTTTGATCAAGTATGCAATATATTAATACTTGATGTACTAGAGTATCAATGTT 8160
Dh 8101 TCTTTGATCAAGTATGCAATATATTAATACTTGATGTACTAGAGTATCAATGTT 8160
Qy 8161 GTTAAGATTTCTTAACTTCAACACCTGCTGCACTTCAATTTGATATCTTTGA 8220
Dh 8161 GTTAAGATTTCTTAACTTCAACACCTGCTGCACTTCAATTTGATATCTTTGA 8220
Qy 8221 CATTAATGTTCTGTAAGTCAACAACTTAATTAAGATTTTGGCTATCAAAATGT 8280
Dh 8221 CATTAATGTTCTGTAAGTCAACAACTTAATTAAGATTTTGGCTATCAAAATGT 8280
Qy 8281 TTAGAGAGCTCTTCTTAAATTTCTTGAAGACAAAGAGATTTCTTTAAGAGATTA 8340
Dh 8281 TTAGAGAGCTCTTCTTAAATTTCTTGAAGACAAAGAGATTTCTTTAAGAGATTA 8340
Qy 8341 TGGTCTAATAAGGAATCATCAGAGATGTTATAGGAAATTAAGAAATCAATATAT 8400
Dh 8341 TGGTCTAATAAGGAATCATCAGAGATGTTATAGGAAATTAAGAAATCAATATAT 8400
Qy 8401 TAGGTGCTCAAGCGCAAGAGATGATGTTGTAATGTAAGCTAAGCATAGACTTA 8460
Dh 8401 TAGGTGCTCAAGCGCAAGAGATGATGTTGTAATGTAAGCTAAGCATAGACTTA 8460
Qy 8461 CTTTTTAAGGTTAATCTGGATCTCAGTGCATCAACAAATCAATCAAAATATAT 8520
Dh 8461 CTTTTTAAGGTTAATCTGGATCTCAGTGCATCAACAAATCAATCAAAATATAT 8520
Qy 8521 TATGTTTGTATGATCTTTTGAAGATGATGATGATGATGATGATGATGATGATGATG 8580
Dh 8521 TATGTTTGTATGATCTTTTGAAGATGATGATGATGATGATGATGATGATGATGATG 8580
Qy 8581 AAATTAATTAAGGATTAATCTGATATCAATGATGATGATGATGATGATGATGATGATG 8640
Dh 8581 AAATTAATTAAGGATTAATCTGATATCAATGATGATGATGATGATGATGATGATGATG 8640
Qy 8641 ATTGTAACACTCAAAATCAATATTAACAAAGTTCTCCATTTTATAGATGACTT 8700
Dh 8641 ATTGTAACACTCAAAATCAATATTAACAAAGTTCTCCATTTTATAGATGACTT 8700
Qy 8701 CTTTTTAATTAAGATTAATCTGATGATGATGATGATGATGATGATGATGATGATG 8760
Dh 8701 CTTTTTAATTAAGATTAATCTGATGATGATGATGATGATGATGATGATGATGATG 8760
Qy 8761 GGAAGATCATGAGACATTAATTTTGTGCAAGCTGCAAGAGAGATCTGTAAT 8820
Dh 8761 GGAAGATCATGAGACATTAATTTTGTGCAAGCTGCAAGAGAGATCTGTAAT 8820
Qy 8821 TTTATGTCAGGTAGGAATGATATGCGCAATTTTAAAGAAAATTAATGATTAAT 8880
Dh 8821 TTTATGTCAGGTAGGAATGATATGCGCAATTTTAAAGAAAATTAATGATTAAT 8880
Qy 8881 ATCTTCTGTTAATCAATTAACAAATCTTCCCTCACTCTTAAATTAATGATTAATG 8940
Dh 8881 ATCTTCTGTTAATCAATTAACAAATCTTCCCTCACTCTTAAATTAATGATTAATG 8940
Qy 8941 ACAATTTTAATCAATGCTGATCAAGAGTTGAACAGTATTTTGGGACAGAGGAGTAG 9000
Dh 8941 ACAATTTTAATCAATGCTGATCAAGAGTTGAACAGTATTTTGGGACAGAGGAGTAG 9000
Qy 9001 TATATATGTTGAGAAATCAATGATGATGATGATGATGATGATGATGATGATGATG 9060
Dh 9001 TATATATGTTGAGAAATCAATGATGATGATGATGATGATGATGATGATGATGATG 9060
Qy 9061 CATGTTCTCTAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 9120
Dh 9061 CATGTTCTCTAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 9120
Qy 9121 ACAATTAATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9180
Dh 9121 ACAATTAATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9180

Qy 9181 TTTATGCAAGAGTACGAGGAGTACTTTCCCTTGACCTATGCTCTTTTTC 9240
 Db 9181 TTTATGCAAGAGTACGAGGAGTACTTTCCCTTGACCTATGCTCTTTTTC 9240
 Qy 9241 TCGTTACATGTTGCTTCTCACAAAAATATATATACCAAGACATGTTCCAAATGAT 9300
 Db 9241 TCGTTACATGTTGCTTCTCACAAAAATATATATACCAAGACATGTTCCAAATGAT 9300
 Qy 9301 ATTATATATTTTGGAGGCTTTTTCACCAACTATATCTTTCATAGTTCTTAAACC 9360
 Db 9301 ATTATATATTTTGGAGGCTTTTTCACCAACTATATCTTTCATAGTTCTTAAACC 9360
 Qy 9361 GTATATATGTTAATCTTACCAAAAACTTATATAGTTTCTCTTAAATACGAGAGTGT 9420
 Db 9361 GTATATATGTTAATCTTACCAAAAACTTATATAGTTTCTCTTAAATACGAGAGTGT 9420
 Qy 9421 TCTTGGAAATGGAATTTGGGTCCTGATTAATCTTTTGGCTGCGCTTCTTCAAGA 9480
 Db 9421 TCTTGGAAATGGAATTTGGGTCCTGATTAATCTTTTGGCTGCGCTTCTTCAAGA 9480
 Qy 9481 TGGACTGGAAGGCTCAATCTCTGCAAGAAAGAAAGAAAGAAATTTCTATTAATG 9540
 Db 9481 TGGACTGGAAGGCTCAATCTCTGCAAGAAAGAAAGAAAGAAATTTCTATTAATG 9540
 Qy 9541 TTGTTAGTTGACACACCCCTAGTTGATCTGATCTGAGCTGTAATCTATTTAGTT 9600
 Db 9541 TTGTTAGTTGACACACCCCTAGTTGATCTGATCTGAGCTGTAATCTATTTAGTT 9600
 Qy 9601 ATCCCATTTATATTTCAATTAACAAATTTGTAATGTTCTGTTGAGCTGTCAGT 9660
 Db 9601 ATCCCATTTATATTTCAATTAACAAATTTGTAATGTTCTGTTGAGCTGTCAGT 9660
 Qy 9661 CATGTAATCTTCTTTTGAATGATATTTGGCTCTATGCTTGCATCAATTAATTAAT 9720
 Db 9661 CATGTAATCTTCTTTTGAATGATATTTGGCTCTATGCTTGCATCAATTAATTAAT 9720
 Qy 9721 GGTGTGTGACAAATGATGATGATGAGGAGGCTTTTGAACCACTTTTCAACAAAT 9780
 Db 9721 GGTGTGTGACAAATGATGATGATGAGGAGGCTTTTGAACCACTTTTCAACAAAT 9780
 Qy 9781 CTATATCTTTTCAACAAATGAACCTTGAATTTCCCTTGAATGAATGACATATCTCT 9840
 Db 9781 CTATATCTTTTCAACAAATGAACCTTGAATTTCCCTTGAATGAATGACATATCTCT 9840
 Qy 9841 GAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9900
 Db 9841 GAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9900
 Qy 9901 CATGCAAAAGTTTATATGTTTATATTTCAATATGAGAACTTGAATTTTCTTACGAC 9960
 Db 9901 CATGCAAAAGTTTATATGTTTATATTTCAATATGAGAACTTGAATTTTCTTACGAC 9960
 Qy 9961 ACATTTCTGGAATATGTCAGTTTGTACTTGAAGGAGCTTATATTTTCTATTTGG 10020
 Db 9961 ACATTTCTGGAATATGTCAGTTTGTACTTGAAGGAGCTTATATTTTCTATTTGG 10020
 Qy 10021 GGGGGGGGGTGAATAGGCGGTTTATACCAATTTGATTTTGAAGAAATCTTATGTCGA 10080
 Db 10021 GGGGGGGGGTGAATAGGCGGTTTATACCAATTTGATTTTGAAGAAATCTTATGTCGA 10080
 Qy 10081 ATTAACTAGGATGATATTTTTCATTAAGGAGCTTTTATGAGCTCACAATGTCACCA 10140
 Db 10081 ATTAACTAGGATGATATTTTTCATTAAGGAGCTTTTATGAGCTCACAATGTCACCA 10140
 Qy 10141 TCAAGGATATCAT 10200
 Db 10141 TCAAGGATATCAT 10200
 Qy 10201 CCAAC 10260
 Db 10201 CCAAC 10260
 Qy 10261 GTCATATCAAGATCAAAACTATGCTTGGCGGAGGAGAAATGAAAAACATGAAGAAATGA 10320

Db 10261 GTCATATCAAGATCAAAACTATGCTTACGCGAGAGAAATGAAAAACATGAAGAAATGA 10320
 Qy 10321 AAAACGCTGACTGACCAATATCTGACATGACGACCAAAATCTGTGACCAACAAAA 10380
 Db 10321 AAAACGCTGACTGACCAATATCTGACATGACGACCAAAATCTGTGACCAACAAAA 10380
 Qy 10381 CTGGAGAAAAAGTCTTAAATCTGAGGCGCTTGAAGAAAGAAAGAGGCTGCAAGATTGC 10440
 Db 10381 CTGGAGAAAAAGTCTTAAATCTGAGGCGCTTGAAGAAAGAAAGAGGCTGCAAGATTGC 10440
 Qy 10441 CATCATGAGATTCACCACTAAGTCAATCTCTGCTTTCATCTGAAATCAATCCG 10500
 Db 10441 CATCATGAGATTCACCACTAAGTCAATCTCTGCTTTCATCTGAAATCAATCCG 10500
 Qy 10501 AGCAAACTCCGAGTAATGCTTTTATTTAGGTTAAGATTCAAAAAATGCAATCATGAG 10560
 Db 10501 AGCAAACTCCGAGTAATGCTTTTATTTAGGTTAAGATTCAAAAAATGCAATCATGAG 10560
 Qy 10561 TTATGACCAATTTAGACGAGCTGAGATTTTATCCAAAGCTGAGACGGGTACTTGA 10620
 Db 10561 TTATGACCAATTTAGACGAGCTGAGATTTTATCCAAAGCTGAGACGGGTACTTGA 10620
 Qy 10621 AGTACCAATCCAAATTTGAATGATCCCACTGCTCAATACAAATAGTTGACATGATGACG 10680
 Db 10621 AGTACCAATCCAAATTTGAATGATCCCACTGCTCAATACAAATAGTTGACATGATGACG 10680
 Qy 10681 GTCCATATGCGAGTAATGACATGAGCGGCGCATGCTGATGATGATGATGATGATGAT 10740
 Db 10681 GTCCATATGCGAGTAATGACATGAGCGGCGCATGCTGATGATGATGATGATGATGAT 10740
 Qy 10741 GCCGTGCGCACCACTGACGAGAGTGTGATGAGGAGGAAAGATGATGCTCCACCAAC 10800
 Db 10741 GCCGTGCGCACCACTGACGAGAGTGTGATGAGGAGGAAAGATGATGCTCCACCAAC 10800
 Qy 10801 ATCCCAAGTTGAAACATTTCTAGAGCCCTTACCATCTCAAAAGGACCAATGATGAC 10860
 Db 10801 ATCCCAAGTTGAAACATTTCTAGAGCCCTTACCATCTCAAAAGGACCAATGATGAC 10860
 Qy 10861 TATCTGATGACAGACCAATTCATGCTGCTCATCTGCTTATCTTGTCTATGACATAC 10920
 Db 10861 TATCTGATGACAGACCAATTCATGCTGCTCATCTGCTTATCTTGTCTATGACATAC 10920
 Qy 10921 TACCTGCAAAAGGCGGATTCAGCCCGACAGCAGCTGCGGCGCGC 10966
 Db 10921 TACCTGCAAAAGGCGGATTCAGCCCGACAGCAGCTGCGGCGCGC 10966

RESULT 4
 AB005788
 LOCUS
 DEFINITION
 AB005788 1895 bp mRNA linear PLN 27-NOV-1999
 Hordeum vulgare mRNA for nicotianamine aminotransferase B, complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.
 REFERENCE
 1 (bases)
 AUTHORS
 Takahashi, M., Yamaguchi, H., Nakanishi, H., Shioiri, T., Nishizawa, N.K. and Mori, S.
 Cloning two genes for nicotianamine aminotransferase, a critical enzyme in iron acquisition (Strategy II) in Gramineous plants
 JOURNAL
 MEDLINE
 20027437
 PUBMED
 10557244
 2 (bases 1 to 1895)
 REFERENCE
 AUTHORS
 Mori, S.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (08-JUL-1997) Satoshi Mori, The University of Tokyo,

Plant Molecular Physiology; Yayoi 1-1-1, Bunkyo-ku, Tokyo 113,
Japan (E-mail: aa1078@hongo.ecc.u-tokyo.ac.jp,
Tel: 03-3812-2111 (ex. 5106), Fax: 03-3812-0544)
Location/Qualifiers

FEATURES

Source

1..1895
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Ehimedaka No.1"
/db_xref="taxon:112509"
/class type="root"
1..1895
/gene="naat-b"
76..1731
/gene="naat-b"
/gene="NAAT-B"
/product="nicotianamine aminotransferase B"
/protein_id="BAAB7053.1"
/db_xref="GI:6469087"
/translation="MATVROSDVAVANGLAVAAAANGKNGHVAALVNGKNGHVD
ADANGKNGHVAADANGKNGHAEATANGHGEATANGKNGHESNGHAEADANG
SNEHEDSANGSNGHAAHAAAEBAEVENPFGAKDQVLAATGAMSTRALYKISA
SVQKGRPVPLAHDPSPVPARTVEADAVAAVTRGQRCYTPAGTGLPARSA
VAEHLQGVPTMSADVFLTAGSTQALVETLIPVLAQTAGANILPRPGVNTBARSA
FNRLVHFDLIPDKGWEIDISLESIAIDKNTAMVINPNDGSSVYSVDHLSKVAE
VAPRKLIVLADEVYGLVGSAPFIPMGVGLTPTLSIGLSKSWIVGMDLKVAA
VADRIKLOETKISTITNYLVNVDTPATFOALPQILENTKEDPFKILIGLKSS
EICYKOIKENKTYICPHKPGSMFVMVKMLHLEIEDDIDPECKLAKESVILCG
SVLGMANVWRTFACVPSLIDGLGRKSKCRKRNKNSDDC"
1895
/gene="naat-b"
/note="15 A nucleotides"

polyA_site

ORIGIN

Query Match 8.9%; Score 971.8; DB 8; Length 1895;
Best Local Similarity 81.7%; Pred. No. 3.8e-115;
Matches 1287; Conservative 0; Mismatches 7; Indels 281; Gaps 3;

QY 579 CGCTACTAGTATTCCTCGTGTAGTCTAGTAGTA--CTCCTCCTCCTCCTCTCTCC 636
DB 3 CGCTACTAGTATTCCTCGTGTAGTCTAGTAGTA--CTCCTCCTCCTCCTCTCTCC 62
QY 637 TACCGTTTCTCATGCGCACCGTACCGCAGACGCGAGTCCGCGGAAACCGGCTTGC 636
DB 63 TACCGTTTCTCATGCGCACCGTACCGCAGACGCGAGTCCGCGGAAACCGGCTTGC 122
QY 697 CGTGGCCGACGCGCGGAAACGCGCAAGACGCGCAATGCGTGTGCTCCCGCGGAAACCG 756
DB 123 CGTGGCCGACGCGCGGAAACGCGCAAGACGCGCAATGCGTGTGCTCCCGCGGAAACCG 182
QY 757 CAAGAGCAACGCGCATGCGTGTAGTCCGACGCGCAACGCGCAAGACGCGCATGCGT 816
DB 183 CAAGAGCAACGCGCATGCGTGTAGTCCGACGCGCAACGCGCAAGACGCGCATGCGT 242
QY 817 GGTGCGGACGCGCAACGCGCAAGACGCGCATGCGGCGCACTGCGAAACCGGCAACG 876
DB 243 GGTGCGGACGCGCAACGCGCAAGACGCGCATGCGGCGCACTGCGAAACCGGCAACG 302
QY 877 CGAGGCAACGCGCAACGCGCAAGACGCGCATGCGGCGCACTGCGAAACCGGCAACG 936
DB 303 CGAGGCAACGCGCAACGCGCAAGACGCGCATGCGGCGCACTGCGAAACCGGCAACG 362
QY 937 CGCGGCAACGCGCAACGCGCAAGACGCGCATGCGGCGCACTGCGAAACCGGCAACG 996
DB 363 CGCGGCAACGCGCAACGCGCAAGACGCGCATGCGGCGCACTGCGAAACCGGCAACG 422
QY 997 CAACGGGCAATGCGGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1056
DB 423 CAACGGGCAATGCGGCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
QY 1057 TGCCAAAG 1116
DB 483 TGCCAAAG 542

QY 1117 CAAGATACGCGGAG 1176
DB 543 CAAGATACGCGGAG 602
QY 1177 GAAACCGGTCGATGTCCTCGGCTTCCGCAACGCGCGTGAAGAGAGAGAGAGAGAGAG 1236
DB 603 GAAACCGGTCGATGTCCTCGGCTTCCGCAACGCGCGTGAAGAGAGAGAGAGAGAGAG 662
QY 1237 CGGCGTGGCAACGCGGCAATGCTCAACGCTGACCCCGCGGCTGCGCTCCCGCGGAG 1296
DB 663 CGGCGTGGCAACGCGGCAATGCTCAACGCTGACCCCGCGGCTGCGCTCCCGCGGAG 722
QY 1297 AAGGTAAACAACAACAACAACAACAACAATTCCTTTGCGGTGTGTCGCGCGGCA 1356
DB 723 A----- 723
QY 1357 ATCCATGATGCGCATGTGCGGCTTTCAGGTGCGGTCCGTCGACCGTTCCTTCT 1416
DB 724 ----- 723
QY 1417 CTTCCCTACGCGCATGAGAAATCTGACCTTCTCCACCTTATACAAACAAACAAAAA 1476
DB 724 ----- 723
QY 1477 ACAAGCGCGGTGAG 1536
DB 724 -----AAGCGGTGAG 779
QY 1537 GGTCTTCTCAACGCGCGGCGGAGCCCAAGCGATGAGTCAATATCCGCTGTCGCCA 1596
DB 780 GGTCTTCTCAACGCGCGGCGGAGCCCAAGCGATGAGTCAATATCCGCTGTCGCCA 839
QY 1597 GACCGCGGCGCAACATTCGTCTCCAGGCGGAGCTACCAATGAGAGCGGCGCG 1656
DB 840 GACCGCGGCGCAACATTCGTCTCCAGGCGGAGCTACCAATGAGAGCGGCGCG 899
QY 1657 CGGTTCAACAGGCTGAGAGTCCGAGATTCGACTCATCCCGAGAGAGGCTGAGAGAT 1716
DB 900 CGGTTCAACAGGCTGAGAGTCCGAGATTCGACTCATCCCGAGAGAGGCTGAGAGAT 959
QY 1717 CGACATGACTCGCTGAATTCATCGCCGCAAGAAACCAACCGCGCATGCTCATTA 1776
DB 960 CGACATGACTCGCTGAATTCATCGCCGCAAGAAACCAACCGCGCATGCTCATTA 1019
QY 1777 CGCCAAACAACCGGTGAGAGAGGTTTACTCTTACGACCATCTGCAAGGTTTCAACATCC 1836
DB 1020 CGCCAAACAACCGGTGAGAGAGGTTTACTCTTACGACCATCTGCAAGGTTTCAACATCC 1066
QY 1837 TTTGCTTGTGTAATGATTCAGTTCAGTGACACTGTGAATTCCTTTTGCCTAACATCCG 1896
DB 1067 ----- 1066
QY 1897 ATATCTGATGATGTTGCTCAATTAAGTCCGCGAGGTGCGGAAAAGCTCGGAATATTGCT 1956
DB 1067 -----AGGTGCGGAGGTGCGGAAAAGCTCGGAATATTGCT 1103
QY 1957 GATTGCGAGAGGTATACGCAAGCTGATTCGAGGACGAGCGCGTTCAATCCCAATGGG 2016
DB 1104 GATTGCGAGAGGTATACGCAAGCTGATTCGAGGACGAGCGCGTTCAATCCCAATGGG 1163
QY 2017 AGTGTGAGCAATCAACCTGTGTCTCATATAGGATCTGTCCAAAGTCATGATAGT 2076
DB 1164 AGTGTGAGCAATCAACCTGTGTCTCATATAGGATCTGTCCAAAGTCATGATAGT 1223
QY 2077 GCTGTGATGCGGCTTGTGATGGGTACCGGTGTCACACCCCAAGAAATCTTAACAGAAAC 2136
DB 1224 GCTGTGATGCGGCTTGTGATGGGTACCGGTGTCACACCCCAAGAAATCTTAACAGAAAC 1283
QY 2137 TAAGTACTTAATC 2151
DB 1284 TAAGTACTTAATC 1298

RESULT 5			
LOCUS	D88273		
DEFINITION	D88273	1660 bp	mRNA
	Hordeum vulgare naat-A	cds.	linear
ACCESSION	D88273		PLN 02-DEC-1999
VERSION	D88273		
KEYWORDS	D88273.2	GI:6498121	
SOURCE	naat-A, nicotianamine aminotransferase A.		
ORGANISM	Hordeum vulgare subsp. vulgare		
	Hordeum vulgare subsp. vulgare		
	Eurycorycia Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
	Pooidaeae; Triticeae; Hordeum.		
REFERENCE	1 (bases)		
AUTHORS	Takahashi,M., Yamaguchi,H., Nakanishi,H., Shiota,T.,		
	Nishizawa,N.K. and Mori,S.		
TITLE	Cloning two genes for nicotianamine aminotransferase, a critical		
	enzyme in iron acquisition (Strategy II) in graminaceous plants		
JOURNAL	Plant Physiol.	121 (3),	947-956 (1999)
MEDLINE	20027437		
PUBMED	10557244		
REFERENCE	2 (bases 1 to 1660)		
AUTHORS	Mori,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-OCT-1996)	Satoshi Mori,	The University of Tokyo,

COMMENT On Dec 1, 1999 this sequence version replaced gl:6469084 Sequence updated (27-Nov-1999).

BOURCE

gene	1. .1660
	/gene="naat-A"
CDS	62. .1447

/gene="naat-A"
 /codon_start=1
 /product="nicotianamine amliotransferase A"
 /protein_id="AA87052.2"
 /db_xref="GI:6496122"
 /translation="MHQNSGHGEBAALAAANGKSGNGHAAANGKSGHAAALAAAEVNER
 ARGQDIALDTGAAGNSIRAIRYKISAVSESGRPVLPPLAHDDPSVFPFAFTVAEABE
 ALAAATGTGFCNCAACVGGPYARASAVAEHLISGVYKADVDLTAGQGLQLEVILIT
 PVLOTAGTANILPRPEYPIYRARAANKLEVHFPGLIDDKMEIDITSLSDIKNTN
 HIAVILSNPNNSGSVSVYHGLAAVAARKLGIIVLADVSGKILGASAPYPMGVGSE
 LIAVILSLSGSVSVYHGLAAVAARKLGIIVLADVSGKILGASAPYPMGVGSE
 EAELKILNTADFPKRIIGLSESESEICREIKENKRIITCHAREGSMFWVKLNTHL
 LILSHIDHIDICCCCLLAARESVILICPSAVIAGMENWATITACVPSLSLQGLERVSPCC
 LNKKRSINGC"

ORIGIN

Query Match	6.8%	Score 744.6;	DB 8;	Length 1660;
Best Local Similarity	80.6%	Pred. No. 5.6e-66;		
Matches 1005; Conservative	0;	Mismatches 4;	Indels 238;	Gaps 2;

[illegible]

OY	6637	GGAGTGGAAATTTCGCCCGGGGCAAGGAGATCTGGCGAGCAACGGGGGCGAAGAACAG	6639
Db	181	GGAATGGAAATTTCCGCCCGGGGCAAGAGACGGCATCTTGGCGACGACGGGGGCGAAGAACAG	240
OY	6697	CATCCGGGCGATACGGTACAAAGATCAGCGGAGCGTGGAGGAGAGCGGGCCGCGCCGT	6756
Db	241	CATCCGGGCGATACGGTACAAAGATCAGCGGAGCGTGGAGGAGAGCGGGCCGCGCCGT	300
OY	6757	GCTGCGGTGGCCCAACGGGTGACCCGTCGTAATCCCGGCTTTCGCGACGGCGGTGAGAGG	6812
Db	301	GCTGCGGTGGCCCAACGGGTGACCCGTCGTAATCCCGGCTTTCGCGACGGCGGTGAGAGG	360
OY	6817	CGAGGACGCGGTCCGCGCGCGCTGCGACCGCGCAAGTTCACTGCTACGCGCGCGGCGT	6876
Db	361	CGAAGACGCGGTCCGCGCGCGCTGCGACCGCGCAAGTTCACTGCTACGCGCGCGGCGT	420
OY	6877	CGGCTCTCCCGCGCGCGAGAGTAACTTACAGCTTCAACGTAATGATGCTGAGCAT	6933
Db	421	CGGCTCTCCCGCGCGCGAGAG	439
OY	6937	GCATGCGCGGTTTACTTACGTGCCCGCGCTGTCTTCCCGGCGCTTCAAAATTTTA	6999
Db	440	-----	439
OY	6997	ACCTTCTATAAGTACTTATAAAAACAACAGCGCGGTAGACAGACACTTGTCAAGGCG	7056
Db	440	-----AGCGCGTAGACAGAGACACTTGTCAAGGCG	469
OY	7057	GTGCCCTTACAAGCTTATGGCGCGGACGAGCGTCTTCTCAACCGCGCGCGGAACTCAAGCGATC	7112
Db	470	GTGCCCTTACAAGCTTATGGCGCGGACGAGCGTCTTCTCAACCGCGCGGGAATCTAAGCGATC	529
OY	7117	GAAATCATAAATCCGCGTGTGCGCCAGACTGCGCGCGCAACATATCTCTTCCCGGAGCA	7176
Db	530	GAAATCATAAATCCGCGTGTGCGCCAGACTGCGCGCGCAACATATCTCTTCCCGGAGCA	589
OY	7177	GCGTATCCAAATTACAGAGCGCGGAGCGGCAATTCAACAGCTGAGAGTCCGCGACTTGCAC	7233
Db	590	GCGTATCCAAATTACAGAGCGCGGAGCGGCAATTCAACAGCTGAGAGTCCGCGACTTGCAC	649
OY	7237	CTCATCCCGGCAAGGGGTGGGAGTGCATTCGACTGCTGGATTCATGCGCGGACAG	7299
Db	650	CTCATCCCGGCAAGGGGTGGGAGTGCATTCGACTGCTGGATTCATGCGCGGACAG	709
OY	7297	AACACCAACCGGATGTCATCATTAACCCAAACAATCCGTCGCGGACGGCTTACTCTAC	7356
Db	710	AACACCAACCGGATGTCATCATTAACCCAAACAATCCGTCGCGGACGGCTTACTCTAC	769
OY	7357	GACCATCTGGCGCAAGTTTGCATTCATGCAATCCTCTGCGCTTGTGATCGACCGGTCTGT	7416
Db	770	GACCATCTGGCGCA	782
OY	7417	TTGAACAATAGATATGAGATTGCGTTTGCTAATCGTGTGCTGATGATGCTGTTGGTTATC	7476
Db	783	-----	782
OY	7477	AGTGTGCGGAGGTGCAAGAAAGCTCCGAAATTTGGTATGCTGTGACGAGGTTTACGGCA	7533
Db	783	AGTGTGCGGAGGTGCAAGAAAGCTCCGAAATTTGGTATGCTGTGACGAGGTTTACGGCA	842
OY	7537	AACGTGTTCTGGGAGAGCGCCCGTTTATCCCGATGGGCGTCTTTGGGCACATTTGCCCGG	7599
Db	843	AACGTGTTCTGGGAGAGCGCCCGTTTATCCCGATGGGCGTCTTTGGGCACATTTGCCCGG	902
OY	7597	TCTTGTCCATTGATCTCTGTCCAAGTGTGGAATGATGCTGAAATGCGCATTTGATGAGG	7656
Db	903	TCTTGTCCATTGATCTCTGTCCAAGTGTGGAATGATGCTGAAATGCGCATTTGATGAGG	962
OY	7657	TGGGGGTGTACAGACCCCAAAAGATTTTAGAGAAAATTAAGATGCT 7703	
Db	963	TGGGGGTGTACAGACCCCAAAAGATTTTAGAGAAAATTAAGATGCT 1009	


```

Db      22958 GTTGAATTTCTTAGAGAGCTCTACCGAATATTTCTTAAGAAATACAGAGGAAGAAATCTTTA 22899
Qy      8331 AGAGGATTTTGGCTCTACTAAAGAAATCATAGAGATTTGTATTATAGGAAATTAAGGAA 8330
Db      22898 AGAGGATTAATGATTTCTTACGAAACATCAGATATTGTCTATAGAGATTAAGGATTA 22839
Qy      8391 ACAATATTAATACGTCTCTCAGAACGAGAGAGATGATGTTGTATGTTAGTAAAGTAA 8450
Db      22838 TTAATATGATCATCTTGCTCTCAGAACGAGAGATCATTTGTATGTTAGTAAAGTAA 22779
Qy      8451 CATAGACTTACTTTTAAAGTTAACTGGAGATCTCAGTATCATTTGTCTCAATCAATCAATC 8510
Db      22778 CTCTTAATCATTTTATCATATAGAGCTGTACT-----AGCATCAATTTATCACTC 22727
Qy      8511 AAAATATTAATATGTTTGTCTATAGATCTTTTGAAGATCATGATTTGAAGATTAATG 8570
Db      22726 AAAAATAATATTATGCTAT--ATTAAATTAGTTGGACCTTAACACTTAATTACAGATC 22669
Qy      8571 AAGAGAGTTGAATTAATTTTAAAGACTAATCTCTGATATCATTTGTCTCAATTTTGT 8630
Db      22668 ATGAGA-----ATTTTGGTGATTTGTGTGAAGAAATATGCTATATTAAGAACT 22617
Qy      8631 ATTAAGTAAATTTGTAACACTCAATCATATTAACAAAGTTCTCTCCATTTTAACTA 8690
Db      22616 TGTATGTTGCATGTTAAATCAATCAATTTTCTTT-----GTGCTCAACATCACTTGT 22564
Qy      8691 AGATGACTTCTCTTTCTATACCATGATTAATTAACCTTCCATGTAACAGTCAACTAACT 8750
Db      22563 ACTTAAGCTTCACTTCCAGTATCTTAATCTTAATTAACAGTGAATTTGAAC 22504
Qy      8751 TACATCTTTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8810
Db      22503 TATATCTTTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22444
Qy      8811 AATCTGTAATTTTATGTCAGATA 8834
Db      22443 AGTCGATGATCTTGTGCCAGATA 22420

RESULT 7
AP005743/c 159075 bp DNA linear PLN 15-JUN-2004
LOCUS      Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
DEFINITION PAC clone: P0543C11.
ACCESSION AP005743 GI:48716873
VERSION    AP005743
KEYWORDS   Oryza sativa (japonica cultivar-group)
SOURCE      Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ORGANISM    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Eriatoidae; Oryzaceae; Oryza.
REFERENCE   1
AUTHORS     Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE       Oryza sativa n1pnpbare (GA3) genomic DNA, chromosome 2, PAC
           clone: P0543C11
JOURNAL     Published Only in Database (2002)
REFERENCE   2 (bases 1 to 159075)
AUTHORS     Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE       Direct Submission
JOURNAL     Submitted (11-SEP-2002) Takuji Sasaki, National Institute of
           Agrobiological Sciences, Rice Genome Research Program, Kanondai
           2-1-2, Tsukuba, Ibaraki 305-8602, Japan
           (e-mail: sasaki@nias.affrc.go.jp, URL: http://rsgp.dna.affrc.go.jp/,
           Tel: 81-298-38-7441, Fax: 81-298-38-7468)
           On Jun 14, 2004 this sequence version replaced gi:38564220.
           Genes were predicted from the integrated results of the following:
           GENSCAN (http://CCR-081.mil.edu/GENSCAN.html), FGENESH
           (http://www.softberry.com/), Genemark.hmm
           (http://opal.biology.gatech.edu/genemark/), Glimmer
           (http://www.cigr.org/cdb/glimmer/glimmer_form.html), RICEHMM
           (http://rsgp.dna.affrc.go.jp/RICEHMM/), SplicePredictor
           (http://bioinformatics.iastate.edu/cgi-bin/bp.cgi), sim4

```

```

FEATURES
Source
gene      4637..4982
           join(<4637..4749,4901..>4982)
           /gene="P0543C11.1"
           /db_xref="GI:48716874"
           /translacion="MTDVTGLTPVMSADSHDSTSHLPKAVEVVGRLIRLCTONK
           GRAGIVFVAPSYDLPKTC"
           5899..7665
           /gene="P0543C11.2"
           5899..7665
           /note="probably inactive due to 5' exon missing in CDS
           pseudogene, GAG-POL precursor"
           /pseudo
           complement(9157..9750)
           /gene="P0543C11.3"
           complement(9157..9750)
           /gene="P0543C11.3"
           /note="polypeptide-like"
           complement(join(10709..10815,11184..11241))
           /gene="P0543C11.4"
           complement(join(10709..10815,11184..11241))
           /note="hypothetical ORF
           predicted by GENSCAN
           this category is not included in IRGSP standard"
           11463..12659
           /gene="P0543C11.5"
           join(<11463..11535,12538..>12659)
           /gene="P0543C11.5"

http://globin.cse.psu.edu/bcm1/docs/sim4.html), gap2
(http://www.cigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI Nonredundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DBJ. Protein homologues of the coding
regions were searched against NCBI Nonredundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from 5' to 3' of the PAC clone.
This sequence of P0543C11 clone has an overlap with P0572A04 clone
(DBJ: AP005009) at 5' end and an overlap with OJNB0026D20 (DBJ:
AP005476) at 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://rsgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1..159075
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0543C11"
4637..4982
/gene="P0543C11.1"
join(<4637..4749,4901..>4982)
/gene="P0543C11.1"
/note="start and end point are not identified"
join(4637..4749,4901..4982)
/gene="P0543C11.1"
/note="predicted by Glimmer etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD23570.1"
/db_xref="GI:48716874"
/translacion="MTDVTGLTPVMSADSHDSTSHLPKAVEVVGRLIRLCTONK
GRAGIVFVAPSYDLPKTC"
5899..7665
/gene="P0543C11.2"
5899..7665
/note="probably inactive due to 5' exon missing in CDS
pseudogene, GAG-POL precursor"
/pseudo
complement(9157..9750)
/gene="P0543C11.3"
complement(9157..9750)
/gene="P0543C11.3"
/note="polypeptide-like"
complement(join(10709..10815,11184..11241))
/gene="P0543C11.4"
complement(join(10709..10815,11184..11241))
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
11463..12659
/gene="P0543C11.5"
join(<11463..11535,12538..>12659)
/gene="P0543C11.5"

```



```

CDS
/note="start and end point are not identified"
join(11463..11535,12538..12659)
/gene="P0543C11.5"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD23571.1"
/db_xref="GI:48716875"
/translation="MSMRHAGVLLGGSIIPACPGSGGLDGVSEEDIVSLPKLR
KTGSGSEVEAYKPAVKOR"
12934..22525
/gene="P0543C11.6"
12934..22525
/gene="P0543C11.6"
/note="probably inactive due to including stop codon(s) in
CDS
pseudogene, polypeptide"
/pseudo
complement(123001..24433)
/gene="P0543C11.7"
complement(join(123001..23488,23511..24433))
/gene="P0543C11.7"
/note="supported by full-length cDNA(s) : AK119914"
complement(123520..23969)
/gene="P0543C11.7"
/note="retrotransposon protein-like"
complement(125886..26575)
/gene="P0543C11.8"
complement(125886..26575)
/gene="P0543C11.8"
/note="probably inactive due to no termination codon in
CDS
pseudogene, orf3 of RIR2"
/pseudo
complement(join(27629..27723,28537..28651))
/gene="P0543C11.9"
complement(join(27629..27723,28537..28651))
/gene="P0543C11.9"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
28948..28562
/gene="P0543C11.10"
join(<28948..28960,29390..29562)
/gene="P0543C11.10"
/note="start and end point are not identified"
join(28948..28960,29390..29562)
/gene="P0543C11.10"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD23572.1"
/db_xref="GI:48716876"
/translation="MTPTAAGRGSPAGAPAAAGCGPVAAGLAARAGAEGRPAWR
LRHVAFFFGELVIRI"
complement(31692..38144)
/gene="P0543C11.11"
complement(join(31692..31946,32033..32113,32444..32602,
32965..33084,33640..33720,33801..33869,34425..34538,
35083..35136,36066..36199,37805..38144))
/gene="P0543C11.11"
/note="supported by full-length cDNA(s) : AK102077"
complement(join(31872..31946,32033..32113,32444..32602,
32965..33084,33640..33720,33801..33869,34425..34538,
35083..35136,36066..36199,37805..38144))
/gene="P0543C11.11"
/note="contains EST(s) : C98634(E0488),A1162173(E0488)
contains full-length cDNA(s) : AK102077"
/codon_start=1
/product="unknown protein"
/protein_id="BAD23573.1"
/db_xref="GI:48716877"
/translation="MAAAEAVALFLHHHLAFGLRISPSVSPRRRSAGVALLA

```

```

VVVVAALLTATTAAGATIESRPADLRSEVDELRLVRLHLESLBENTKLSKAN
LEENSLNIGMEHDIILNNKYESTKSSKSPENRYALREVLMLVVRKMN
ADISIELANGAEKRVSLISFYKMEGVIAEMIOIROLEAVLTMKXHSR
LSEITAYKMGKIDLVKTFNMLGTPLMGYSYTSKSHYGRSTIQANRPYHVS
RFCKAICGHIKRDVKNVFLGGSISRCSISAPYQKLFMLLAQHFHKVQIFLD
AMRSNYSRGPATIEITPCLAYFVVISPMWILFWLSTRGSKK"
complement(39218..39454)
/gene="P0543C11.12"
complement(39218..39454)
/gene="P0543C11.12"
/note="hypothetical ORF
predicted by Glimmer"
this category is not included in IRGSP standard"
40782..46296
/gene="P0543C11.13"
join(40782..41047,44517..44580,45626..45698,45821..45876,
45968..46296)
/gene="P0543C11.13"
/note="supported by full-length cDNA(s) : AK066730"
join(44525..44580,45626..45698,45821..45876,45968..46094)
/gene="P0543C11.13"
/note="contains EST(s) : AU031536(IE61810),AU031537(IE61810)

Query Match 3.3%; Score 366.8; DB 8; Length 159075;
Best Local Similarity 55.7%; Pred. No. 2.3e-37;
Matches 1082; Conservative 0; Mismatches 692; Indels 170; Gaps 13;

CDS
7008 GTACCTTATTAACAAACAGCCCGCTAGAGAGCACTGTCTCAGGCGCTGACAA 7067
96676 GTCCCTTTCTTGAATATAGTCTGTGAGATCAATTGTCTACGCACTCCCATCA 96617
7068 GCTATCGCCGACGAGTCTTCTCAACCGCGCGGAACTCAGCGCATCAATCAAT 7127
96616 GCTATCTTGAGATCATCTTCTTAACCGTGAGAACTCAGCGCATCAATCAAT 96557
7128 CCCGTCGCGCCAGACTGCGCGCCCAATATGCTTCCCGCGCAGCTATCCAA 7187
96556 CTCATCTCTTGCCCAACT---GGCAAAACATATTGCTTCTTAGACAGGCTAACCA 96500
7188 TTACGAGGCGCGGCAATCAACAGCTGAGGTCGCGGCACTTGCACTCAATCCCG 7247
96499 CTATGAACTCGAGCGCGCTTCAACCACTTGAAGTTCGCTTGATCTTATCTGA 96440
7248 CAAAGGTCGAGANTGACATCGACTGCTGATCATCGCCAGACCAACCAACCG 7307
96439 GAAAGCTGGAGATTGACCTTAACCTCCAGATCTATTGGGACAAAGAACTACTGC 96380
7308 GATGTCATCATTAACCAACCAATCCGTCGCGCAGCTTACTCTTACCACTCTGC 7367
96379 GATAGTCATCATTAATCCCAATATCATGCGGAAATGTGTACCTTACGAGCATTTATC 96320
7368 CAAGTTTTCATCCATGATCCTCTGCGCTGATTCGACCGGTCGTTGAACATAGT 7427
96319 CAAGTTCAGCTGTTATTTAATTAATTAATTAATTAATTAATTAATTAATTA 96260
7428 ATATGATTCGCTTCTATCTGTCTGATGATGCTGTTTGTATATCAAGTCGGAG 7487
96259 TTA-----ATGTGATATGATTAATTTATTAATTAATTAATTAATTAATTA 96221
7488 GTGCGAAGAGTCTCGAATATTGATGATGCTGATGAGAGTTTACGCAAACTGTTCTG 7547
96220 GTAGCAAGAAAGTTGGATATTGTAATTAATTAATTAATTAATTAATTAATTAAT 96161
7548 GGCAGCGCCCGTTATCCGATGCGGCTTTGGGACATTCGCCGCTGTGCATT 7607
96160 GGGAGTTCCTCATTTGTCCTCAATGGTGTCTTTGGGACATCTACCAATATTACCAT 96101
7608 GGAATCTCTGCAAGTCGTGATGATGCTGTGATGCGCACTTGATGATGATGCGGTGAT 7667
96100 GGAATCTCTGCAAGTCGTGATGATGCTGTGATGCGCACTTGATGATGATGCGGTGAT 96041
7668 GACCCCAAAAGTTTAAAGAAATAAGTATGCTT-----TAGCTCCCATCA 7717

```


Dh	96040	GACCCCAAGAAAGACTACACAAAGAACCAAGTAAATATATATATTAATCTGTTTGGAA	95991
Qy	7718	TTCTTTCATATGCTACTGTGGGATTAAGTATTTTGGCTAAATTTGACTGCTTGT	7777
Dh	95980	CTTTTGTGMAATGTAATATTTGAAATATATATTTTCTAAAGTATACTAC---TTTCT	959244
Qy	7778	ATTGAGATCTCTACGTCTATTACGAAATTTACTTATATGTCTCAACGAGACCAGAACCTTC	7837
Dh	95923	TTGTAAGTTGCAACATTAAATTAATCTTAAATTTCTTAATGTTTCAACTATCTACGCACTTTC	95864
Qy	7838	GTTCAAGTTAGCTTGTGTTCTTGCC-----TATTTGCTCATGT	7878
Dh	95863	ATTGAGGTTAGTGTTTTTCCTTAAACAGCCAGATGTTGGGTTATATATTTGGTCATTT	95804
Qy	7879	CCCGTGTCGATGTCAAATAGCCGGCTCAAGTTAGTATATAGATTTTGTAAAGCT	7938
Dh	95803	ATCTGCTTATATGTAATGTTTGGTGACATATGTTGTGTGCTGCTTCGAAATTTGATG	957444
Qy	7939	GAATGTGGAAGTCCACATGATGGAAGAAAGATACATCTATTTTATGTCATTTCCCTTGG	7998
Dh	95743	GCTGTCAAAATGACAGTGTGGTGCTACATGTTGTTGAGTGTGAAAACCTACTGGT	95684
Qy	7999	TT-----TGTTGATTTCCATTAATAATTAATTAACAAAGCCGAAACCAATTTGAATA	8052
Dh	95683	TTAAATATATCTATCGAATCTTACCAAAATAAATAGTACTATGAAATCAGAGTTTAAAC	95624
Qy	8053	GAACATATTTTCTTGAATAATATACA-----TTGTATTTTGAAGATGC	8095
Dh	95623	AAACATAGTTTCATCAAAAGTACAAAGAAAGTTTATATCAGTTTATATGTTCAATTTCAAT	95564
Qy	8096	CATATTTCTTTTCGATCAAGTATGCAATATATTAATACTTGATTTACTACAGATATACC	8155
Dh	95563	CCATATTTCTGTAAACATATATATATGTTTAAAAATTAATTGAGAAATGAGGTTGT	95504
Qy	8156	ATGTGT-----	8162
Dh	95503	ATTTTGTTCCTACCTGAGGTGGTGAATTTTAAGTATATGTGAAGTTCAATGTTAC	954444
Qy	8163	-----TAGAATTTCTTAACTTAACAACACCTGTCTCGCATCTTCAATTTTGG	8210
Dh	95443	GTTAAAGTAAAGCATACAAATCTTATATCTGATGATTAATTTATGGTATCTTGCTTAAAT	95384
Qy	8211	ATATCTTGACATATATGTTCTCTATGATTTCAACAACATTAATTTATGATTTTGTGT	8270
Dh	95383	ATTATATTTGCGCTCGTGATATCAATTTTATCTCAAAATTAATGAATGTGTGATGTGT--TGTT	95325
Qy	8271	ATCAATTTGTTTGAAGAGCTCTTCTCTAAATTTCTTGAGAAACAAAGACAGATTTCTTTA	8330
Dh	95324	GTTGAATTTCTTGAAGAGCTCTTACCGAATTTCTTAAAGATATCCAAAGAAAGATTTCTTTA	95285
Qy	8331	AGAGAGTTATTTGCTATCTAAAGAAATCATCAGAGATATGTTATGAGAAATTAAGAAA	8390
Dh	95264	AGAGAGTAAATTTGATTTGCTTACGGAACAATCAATATTTTGCTATAGAGAAATTAAGATA	95205
Qy	8391	ACAAATATTTAAGTGCTCTCAAGAGCGAAGAGATCGATGTTTGAATAGTAAAGTAAAG	8450
Dh	95204	TTAATATCATCACTGTCTCTCAAGCGCGAAGAGATCAAGTTTGAATGTAAGATTTTC	95145
Qy	8451	CATAGACTTACTTTTAAGGTTATCTGGAGTCTCAGTGATCCAAACAAATCAATCAATC	8510
Dh	95144	CTCTTAATCATTTTATCATTTAGAGCTGTACT-----AGATCAATTTTATCACATCTC	95093
Qy	8511	AAAAATATATATGTTTGTCTATGATATCTTTTGAAGATGATGATTTGAAGATTAATG	8570
Dh	95092	AAAAAATATATATGCTAT--ATTATATTTAGTTGGACCTTAAACCTTAATTTCAAGAGTC	95035
Qy	8571	AAGAGAGTTGAATATATTTTGAAGACTAATCTTCTGATATCATTTGTTCCATTTTGTGT	8630
Dh	95034	ATGAGA-----AATTTTGGATATGTGTGTAAGATATATGCTATATATTAAGAAAGT	94983
Qy	8631	ATTATCTGTAATTTGTATACACTCAAAATCAATTAACAAAAAGTTTCTCCCATTTTATGTA	8690
Dh	94982	TGTAGTTTGCATGTTAAATCAATTTTCTTTT-----GTGTCAACAACTACTGTGC	94930

QY		8691	AGATTGACCTTCCTTTATATAACAGTAAATTAACCTTCATGCATGTAAACAGGTCAACTAAT	8750
DB		94299	ACTTAGCTTACTCTCCAGTATATCTTACATTTGAATCTATTAACAAGTGAAATTTGAACC	9487
QY		8751	TACATCTTTTGAGAGAGATCCATGACGACATMAATTTTTGGCTGCGAAGCTCGCAAGAAG	8810
DB		94869	TATATCTTTTGAGAGGAATCCATGATGATGTGATTTTGTGCGCAACTTCGAAAAGAG	9481
QY		8811	AATCTGTAATTTATATGCCAGTA	8834
DB		94809	AGTGGTGATTTCTTGCCAGSTA	94786
RESULT 8				
LOCUS	BT009504		1035 bp	mRNA
DEFINITION	Triticum aestivum clone wrl.pk0085.h9.fls		full insert mRNA	
ACCESSION	BT009504			
VERSION	BT009504.1		GI:32129055	
KEYWORDS	FLU CDNA.			
SOURCE	Triticum aestivum (bread wheat)			
ORGANISM	Triticum aestivum			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae;			
JOURNAL	Poideae; Triticeae; Triticum.			
FEATURES	1 (bases 1 to 1035)			
source	1..1035			
location/Qualifiers	/organism="Triticum aestivum"			
	/mol_type="mRNA"			
	/db_xref="taxon:4565"			
	/clone="wrl.pk0085.h9.fls"			
ORIGIN				
Query Match	3.0%; Score 326; DB 8; Length 1035;			
Best Local Similarity	75.5%; Pred. No. 3..1e-33;			
Matches	465; Conservative 0; Mismatches 55; Indels 96; Gaps 1;			
QY	1554	GCGGAGCCGAGGCGCATGAGGTCAATATCCCGGTGCTGGGCCAGACCGCGGCCCAACA	16133	
DB	1	GCAAGAGCCGCAAGCATTTGAGGTCAATATCCAGTTCTGGGCCCAAACCGCGGCCCAACA	60	
QY	1614	TTTCGCTCCCCAGGCGCGTCAACCAATACGAGGGGCGCGCGGTTCACAAAGGCTGG	16733	
DB	61	TTTCGCTCCCCAGGCGCGTCAACCAATACGAGGGGCGCGCGGTTCACAAAGGCTGG	120	
QY	1674	AGGTCCGCGCATTTTGACCTCATCCCAGCAAGAGGGGTGGAGATGACATGCACTGCTGG	17333	
DB	121	AGGTTCGACATTTTGACCTCATCCCAGGAGAGGGGTGGAGATGACATGCACTGCTGG	180	
QY	1734	AATTCATGCGCGCAGCAAGAACACACCGCGCATGTCTCATATTAACCCAACAACCGTGGC	17933	
DB	181	AATTCATGCGCGCAGCAAGAACACACCTCGATGTCTCATATTAACCCAACAACCTCATGTG	240	
QY	1794	GCAGCGTTTACTCTTACGACATGTGCAAGGTTTTCATCATCTTTGCTGCTGATATAT	18533	
DB	241	GCAGCGTTTACTCTTACGACATTTTACCAAGTT-----	275	
QY	1854	GGATTCACTTCAAGTCACTGCTGATATCTTTTGGCAATGCACTACTGATGATGTC	19133	
DB	276	-----	275	
QY	1914	TCAATTAGTGCAGGAGTGGCGAAAGGCTCGGAATATTTGGTATGTTGTCAGAGGTAT	19733	
DB	276	-----GCAAGAGTGGCAAAGAGCTCGGAATATTTGGTATGTTGTCAGAGGTAT	324	

The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0477B05 clone has an overlap with O52NB0010K08 (DBJ: AP005532) clone at 5' end and with P0572A04 (DBJ: AP005009) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://xrp.dna.affrc.go.jp/Genomeseg.html>.

location/Qualifiers

1. 14216

 organism="Oryza sativa (japonica cultivar-group)"

 mol_type="genomic DNA"

 cultivar="Nipponbare"

 db_xref="taxon:39947"

 chromosome="2"

 clone="P0477B05"

 439..5934

 /gene="P0477B05.1"

 join(<439..549,1811..1882,2141..2230,2346..2388,2631..2737,4724..4819,5249..5439,5539..5608,5827..>5934)

 /gene="P0477B05.1"

 /note="start and end point are not identified"

 join(439..549,1811..1882,2141..2230,2346..2388,2631..2737,4724..4819,5249..5439,5539..5608,5827..5934)

 /gene="P0477B05.1"

 /codon_start=1

 /product="putative cell differentiation protein Rcdj"

 /protein_id="BAD3249.1"

 /db_xref="GI:48716579"

 /translacion="MSTIQVNPPTDPELAQRVLVDLDPDLRGHALAEKRRKEMFPDLAPLIRSGFRIIVLOEIIYVYVPLSPPTSLIASSVCNALALOCASHEPESHFLAHLPIVCSLEPNTSKTPREPIRLTSGVIGLAKETGVNLLQCFVPLCLHMAVSESLKTYATPIIEKIYLDAGLGTCATADRFPAVGTALAGVTSNDKPSRLRLKIHICLYLRITDNPREGALQCLPTLLIDETFNNTLKDDPTMQWDELIVKVRSGSGGLPPPVQHMPRV"

 complement(join(8981..9460,9615..9812,9980..11942,13202..14057,14312..14637))

 /gene="P0477B05.2-1"

 complement(join(6981..9460,9615..9812,9980..11942,13202..14057,14312..14637))

 /gene="P0477B05.2-1"

 /note="supported by full-length cDNA(s) : AK067818"

 complement(join(9218..9460,9615..9812,9980..11942,13202..14016))

 /gene="P0477B05.2-1"

 /note="contains full-length cDNA(s) : AK067818"

 /codon_start=1

 /product="putative LZ-NBS-LRR class RGA"

 /protein_id="BAD3250.1"

 /db_xref="GI:48716580"

 /translacion="MAVVAGMPTLILAKLTLLMDQVLCIOLNSARGDIIPIKAEISIOALEKIFENVOTDROVKIWRDVRQEGYDMDIINDPNVHETHLLARHELKPKPKGSLSLLRRAMRYRYATDIKAIIRLVNETSARDKRYGTIVAAATETAIIDRLIGTYGEATKLGISGPELTYLLMDSKNSKNKIKYISVGVGKGTTLNAVITYOQLGQFECHAFVSALKPDILKVLSSIRQIEBOGYAMETWAOEIKIRDEIKERKYLIVIDDIWESAWECIECALIENDRSGRIITTSVLDVAATPCCSEBVDIVYKLOPISNRSKILFYKIFCYCEDGCPLEKDISEKTRKCGVPLAITIISLLAMRPONLQWMBVHNIJGSGLEKSHVENMHILISYVDLPAHRAQFLVLSYIPEDVNIQDOLIRWISYEGTIEDMDTLIEGGYRNELIRMSIOPAYIDSHGVACRYHVDVLDITLSNE TNVTSISGGQPTTYHPRARRCLQNSITYDTHIOEKMKRVSRLIFPHATNLSLSFRLLRLVLDLEGQDOLKNQIIEGIDLPRLSLVLDNTNIGSPKIKIGNSCHTLIDHHTITELPSTVYHRLVRLIDASVKLPDGIKNGECQOEISLVGISKSPNFKELGSLTEIRLQIIESGTAMHDSYENTLIDSLCNHAKICDLVHIGCKLSTEPISNRS PRVRLVYSGQLSLIPWINSLSLTIDILINVLVDDDDRLSGALOSICLRLNFIKIREBLVGTETIAKPHMLAEFSFTTNAMGILFQSYSPRLNLELALANVREKTDI GHEHLSLNDVTARIDCRDSSIFEVQNAADAIRAIRAIVNSNQPKHYIRHYHMLDE VKIQRKTEEBEKILENRIIPWGGVATCDIKVAPKRLSITTCGIIITDALFSTLD KQGERHTLSMGIGSGVOLIDDERFLFWEIVGTGFENLSEATISITLVTNVA SYGPGQCGTPTFRTPRKNSCIVGFGFSRYGLDAVVYFHFM"

 complement(10700..15356)

 /gene="P0477B05.2-2"

 complement(10700..15356)

 /gene="P0477B05.2-2"

 /note="contains full-length cDNA(s) : AK101276"

 non-coding transcript

FEATURES

Source

gene

mRNA

CDS

gene

mRNA

CDS

misc_feature

probably inactive due to including stop codon(s) in CDS"

gene
15125..15199
/gene="P0477B05.3"
/note="start and end point are not identified"

mRNA
15125..15199
/gene="P0477B05.3"
/note="start and end point are not identified"

CDS
15125..15199
/gene="P0477B05.3"
/note="predicted by GENSCAN etc."

gene
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD3251.1"
/db_xref="GI:48716581"
/translation="MADGSPMKQDAPEPTERADRGRLRGSSLEGGRTRIRING
GRSTASIVGFRRRDAGCGRGQSRFRREVQMGIFLTGTRGDDRTVWRMR
LAKSGARSTVTLPLPTAIG"
complement(16879..17461)
/gene="P0477B05.4"
complement(join(<16879..17026,17073..17461))
/note="start and end point are not identified"
complement(join(16879..17026,17073..17461))
/gene="P0477B05.4"
/note="predicted by GENSCAN etc."

CDS
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD3252.1"
/db_xref="GI:48716582"
/translation="MGAMAMAGSRGSRARRRRAVRWRHRRARGRKELMWADGRRR
PMHDDGVVATRPDGRVRLQPLGPAIGAREGGGLVGARPGAAARSRKSSATTA
SVAFQGAKEAVGEGEDRREREGPEDRRSPADAKGSDVAAVAEVAADVDRGGR
QCGRRRQWRARECRERI"
18490..23166
/gene="P0477B05.5"
join(<18490..18600,22594..23166)
/gene="P0477B05.5"
/note="start and end point are not identified"
join(18490..18600,22594..23166)
/gene="P0477B05.5"
/note="predicted by Genemark.hmm etc."

gene
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD3253.1"
/db_xref="GI:48716583"
/translation="MEDHVSPESEAKAREKGVGDGCVGDDGVSSATRELYCTISG
GEQDAATPSLRGDAPQGRRAASCGRRPQVLAAGVLRGAARPAALARRGG
RRRAEAGHGGRRAAARPRGRPRRAEALRGAGGGRHRRARVERRRRV
APPRGGERRLAHAVAIEGRRSPCSSVAPVPVPGCPGDAEMEFGVPTIG
QPNRAKEG"
23948..24894
/gene="P0477B05.6"
join(<23948..24234,24323..24377,24721..24894)
/gene="P0477B05.6"
/note="start and end point are not identified"
join(23948..24234,24323..24377,24721..24894)
/gene="P0477B05.6"
/note="contains EST(s) : C72041(E0847)"

gene
/codon_start=1
/product="unknown protein"
/protein_id="BAD3254.1"
/db_xref="GI:48716584"
/translation="MENGSGDPAASAAVRFGAANPAMEARSOSTRALVYVYCTLDG
DARSVAPLGHDPAPACFRPAATGAVVAAAAGAHNSAAPALAEKCRSTKVA
QAOVTYGPSTQGRRRQAEQMRASAAEMRRARFRAEEMAAAGRVTSIDVMSRLA
GELSPAVANLT"
26082..26672
/gene="P0477B05.7"
join(26002..26082,26672..27153,27278..27827)
/gene="P0477B05.7"
/note="retroelement-like"
complement(29206..32244)
/gene="P0477B05.8"
complement(29206..32244)

gene
/note="retroelement-like"
complement(29206..32244)
/gene="P0477B05.8"
complement(29206..32244)

/gene="P0477B05.8"

Query Match 2.5%; Score 273.8; DB 8; Length 142166;
Best Local Similarity 65.9%; Pred. No. 2e-25;
Matches 446; Conservative 0; Mismatches 222; Indels 9; Gaps 3;

Qy 1480 CAGCGCGGTGAGAGACCTGTGCGAGGGGGTGCCTATCATCGGCGGACGCT 1539
Db 40983 CAGCGCGGTGAGAGACCTGTGCGAGGGGGTGCCTATCATCGGCGGACGCT 41042

Qy 1540 CTTCCTACCGCGCGCGGAGCCAGCGGATGAGGTATATACCGGTGTGCGGAC 1599
Db 41043 GATGCTACCGCGCGGTGAGAGCCAGCGGTGAGATATATATGTCGCTGCGG 41099

Qy 1600 CGCGGCGGCGGACATTTGCTTCCCGAGCCAGGTATCCAACTATCGAGGGCGCGCGC 1659
Db 41100 GCGCGGCGGCGGACATTTGCTTCCCGAGCCAGGTATCCAACTATCGAGGGCGCGCGC 41159

Qy 1660 GTTCAACAGGCTGAGAGGTCCGCGATTTGACCTCATCCCGAGAGGGGTGGAGATCGA 1719
Db 41160 CTGAGCGGCTGAGAGGTCCGCGATTTGACCTCATCCCGAGAGGGGTGGAGATCGA 41219

Qy 1720 CATGACTGCTGAGATCATTCGCGCAAGAACACCGCGCATGTATCAATAAACC 1779
Db 41220 CTTGCGCGGCGTGAAGGCGCTCGCGAGCGCAACCGGTGCGATTCGTAACCC 41279

Qy 1780 CACACACCGTGGCGGAGCGTTTACTCTACGACATCTGTCCAGGTTTCAATCCTTT 1839
Db 41280 CACACACCGTGGCGGAGCGTTTACTCTACGACATCTGTCCAGGTTTGAATTTA 41339

Qy 1840 GCGTGTGTA--ATAGATTTCACTGAGTACGACCTGCTGAATCTTTTGGCAATGCA 1897
Db 41340 ATACATCTGATCAATATATATTCGATGCGAGATGATACAAATGTATATACCGCA 41399

Qy 1898 TACTGACTGATGTTGCTCAATTTAGTGCAGAGGTGCGAAAGGCTCGAATATTGCTG 1957
Db 41400 TTAAT----ATGCTAGTACTGATGATGCGAGAGCGCGAGAGTGGGATATGCTG 41455

Qy 1958 ATTGCTACGAGGTATACGCGAGAGCTGTTCTGGGCGGCGCGCTTCATCCCAATGGA 2017
Db 41456 ATCAGCAGCAGAGGTATACGCGAGAGCTGTTCTGGGCGGCGCGCTTCATCCCAATGGA 41515

Qy 2018 GTGTTGGGCAATCACCCGTCGTCATATAGGCTGCTGTCATGATGATGATG 2077
Db 41516 GTGTTGGGCAATCACCCGTCGTCATATAGGCTGCTGTCATGATGATGATG 41575

Qy 2078 CTTGATGCGGCTGTGATGAGGTATGCGGTGATACGACCCAGAAAGATCTTACAGAAACT 2137
Db 41576 CCGGCTGCGGCGCTGCGCTGATGCGCGGATGCGGATGCGGATGCTCAGGAAACAG 41635

Qy 2138 AAGGTACTTAATCTCT 2154
Db 41636 AAGGTACTTAATCTCT 41652

RESULT 10
AP005532

LOCUS 184760 bp DNA linear PLN 15-JUN-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone:OSJNBa0010K08.
ACCESSION AP005532
VERSION AP005532.3 GI:48716663
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehnhartoidae; Oryzaceae; Oryza.

REFERENCE
1 Sasaki,T., Matsumoto,T. and Katayose,Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OSJNBa0010K08
JOURNAL Published only in Database (2002)


```

mRNA      /gene="OSJNBa0010K08.8-2"
          <28390..>28918
          /gene="OSJNBa0010K08.8-2"
          /note="supported by full-length cDNA(s) : AK063049"
misc_feature
          /gene="OSJNBa0010K08.8-2"
          /note="contains full-length cDNA(s) : AK063049
          non-coding transcript
          probably inactive due to including stop codon(s) in CDS"
          complement(join(34424..34584,34767..34818,34903..35051,
          36403..36567,36919..37030,37086..37128,37297..37370))
          /gene="OSJNBa0010K08.9"
          complement(join(34424..34584,34767..34818,34903..35051,
          36403..36567,36919..37030,37086..37128,37297..37370))
          /gene="OSJNBa0010K08.9"
          /note="hypothetical ORF
          predicted by GENSCAN
          this category is not included in IRGSP standard"
gene      /gene="OSJNBa0010K08.10"
          complement(37645..41835)
          /gene="OSJNBa0010K08.10"
          complement(join(37645..38343,39677..39796,40494..40872,
          41186..41222,41582..41831))
          /gene="OSJNBa0010K08.10"
          /note="supported by full-length cDNA(s) : AK100174"
          complement(join(37750..38343,39677..39796,40494..40872,
          41186..41222,41582..41834))
          /gene="OSJNBa0010K08.10"
          /note="supported by full-length cDNA(s) : AK104411"
          complement(join(38143..38343,39677..39796,40494..40872,
          41186..41222,41582..41624))
mRNA      /gene="OSJNBa0010K08.10"
          complement(join(37756..38343,39677..39796,40494..40872,
          41186..41222,41582..41834))
          /gene="OSJNBa0010K08.10"
          /note="supported by full-length cDNA(s) : AK104411"
          complement(join(38143..38343,39677..39796,40494..40872,
          41186..41222,41582..41624))
CDS
Query Match 2.5%; Score 273.8; DB 8; Length 184760;
Best Local Similarity 65.9%; Pred. No. 2e-25;
Matches 446; Conservative 0; Mismatches 222; Indels 9; Gaps 3;
QY 1480 CAGGCGCGTGGCAGACCTGTGCGAGGGGTCGCTACATGCTATGCGCGACGACT 1539
DB 179382 CAGGCGGCTGGGCGGACCTGTGCGAGAGCTCCCTACCGCGCTGCGCGGAGCT 179441
QY 1540 CTTCCTCAACCGCGCGGAGACCCAGGCGATGAGTCAATATCCGCTGTGCGCCAGAC 1539
DB 179442 GGTGTCAACCGCGCGCTGCAACGCGGTGAGATGATGTCGTGCTCGC---GTC 179498
QY 1600 CGCGCGCGCCCAATCTTGTCTCCAGGCGAGCTACCCAAATCTAGAGCGCGCGCGC 1659
DB 179499 GCCGCGCGCCCAACGTGTCTGCCGCGCGGCTACCCGCTGTACGCGTCCGCGCGC 179558
QY 1660 GTTCAACAGCTGAGAGTCCGCGCATTTGACCTATCCCGCAACAGGGGTGGAGATGGA 1719
DB 179559 CCTGAGCGGCTTCAGAGTTCGCGCACTTGACCTCTCCCGACAGCGAGTGGAGTGA 179618
QY 1720 CATGCACTCGCTGATCATCGCGAGACAGAACACACCGCATGTGATCATTAACCC 1779
DB 179619 CCTGCGCGGCGTCAAGGCGCTCGCGAGCAACACCGTCCGATGTCATCGTCAACC 179678
QY 1780 CAACAACCGCTGGCGAGCGTTTACTCTAGACCATCTGTCCAAAGTTTCAATCCTTT 1839
DB 179679 CAACAACCGCTGGCGCTGCTACTCCCGGACCACTCCCAAGTTTGTGATTGATTA 179738
QY 1840 GCCTTGCTGA--ATAAGATTCAAGTTAGTGACCTGTGAAATCTTTTGCATGCGCA 1897
DB 179739 ATACATCTGATCAATCAATTCGATCGCGATGATGCAATCTTAATACATACCGGA 179798
QY 1898 TACTGATGATGTGCTCAATTAAGTCCGAGAGGTGGCGAAAGGCTCGAATATTGAGTG 1957

```

```

DB 17979 TTAAT---ATCTGATAGTACTAGATCGACAGCGGAGGAAGTGGGATAGGTG 179854
QY 1558 ATTGCTACAGGATATACGCAAGCTGTTTGGCGAGCGCCCGTTTATCTCCATGGGA 2017
DB 179855 ATCAGCGACGAGGGTACGACACTTCCGCTTCCGAGCAAGCCGTTGTCGATGGG 179914
QY 2018 GTGTTTGGGCATCATCACCCCTGTGCTGTCATAGAGTCTGTGTCAGATGATAGTG 2077
DB 179915 GTGTTTGGCGACGTCGCGCGCGCTGATGACGTCGGCGCGCATCTCCAGCGTGTG 179974
QY 2078 CTTGATGCGCGCTTGAATGGGTAGCGGTGACACCCAGCAAGATCTTACAGAACT 2137
DB 179975 CCGGCTGCGCGCTTCGCGTGTGATCGCGCGACCGATCCCAAGATCTCAGAAACAG 180034
QY 2138 AAGTACTTAATCTCT 2154
DB 180035 AAGGTCTTAATTGCT 180051

RESULT 11
AP004679 111173 bp DNA linear PLN 24-JUN-2004
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,
BAC clone:OSJNBa0021N09, complete sequence.
ACCESSION
AP004679
VERSION
AP004679.2 GI:49169756
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryote; Vascularplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharitoidae; Oryzaceae; Oryza.
REFERENCE
1 Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare (Ga3) genomic DNA, chromosome 6, BAC
clone:OSJNBa0021N09
Published only in Database (2002)
2 (bases 1 to 111173)
JOURNAL
Oryza sativa (japonica cultivar-group)
AUTHORS
Sasaki,T., Matsumoto,T. and Yamamoto,K.
DIRECT SUBMISSION
Submitted (30-JUN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannonnai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jun 23, 2004 this sequence version replaced gi:18447938.
The orientation of the sequence is from -21m3 to M13rev of the BAC
clone.
FEATURES
location/Qualifiers
1..111173
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="6"
/clone="OSJNBa0021N09"
ORIGIN
Query Match 1.9%; Score 213; DB 8; Length 111173;
Best Local Similarity 60.4%; Pred. No. 1.3e-17;
Matches 408; Conservative 0; Mismatches 255; Indels 12; Gaps 3;
QY 1475 AAACACGCGCGCTGGAGACACCTTGTGCAAGGCGCTGCCATCATGCTATCGGCCGAC 1534
DB 41525 AAATCCAGACTATTTGAGAGTACTTCTGTGATCTTCTTCAAGCTTTTGACAGAT 41584
QY 1535 GAGCTTCTCTACCGCGCGGAGACCCAGCGATGAGGTATTAATCCGCTGTGCGCC 1594
DB 41585 GATTTTCTCTACATCTGAGAGTACCAAGCAATGATGTATGTCTGTTTGGC 41644
QY 1595 CAGACGCGCGCGCAATCTTGTCTCCAGGCGCAGCTACCCAAATCTAGAGCGCGC 1654
DB 41645 CA---ACGAGGTGCAATATATTGCTTCAAAAGCCGCGTACCAAAACATGAAGCAT 41701

```

Oy	1655	GC CGCGTTTCAACAGGCTGAGAGGTC	CGCGCAATTTTGACCTCATCCCGACAAAGGCGTGGAG	1714
Db	41702	CGCGTGTTCACAGAGATGAGAGTCCG	CTTATGATCTTTGTCAGAGAGATGGAG	41761
Oy	1715	ATGCACATTCGACTGCGCTGGAAATC	ATCGCCGACAGAACACCGCCATGGTCATCAT	1774
Db	41762	ATTATATGTGTAAGCTGTGTAAGCTT	TAGCAGATGAGAAATATCTGTTCCAATAGATTA	41821
Oy	1775	AACCCCAACAACCGGTGCGGACGGCT	TACTCTTACAGACATCTGTCCCAAGG--	1831
Db	41822	AACCCCAATAACCTTTGTGTATATGT	TAATACCTTATAGAGATCTGTCCAAAGTTGGT	41881
Oy	1832	CATCTTTGCTTGCTGTAATATGGAAT	CACTTCACTGACACTGCTGTAATCTTTTGGCA	1891
Db	41882	ATGCTGTACAACTGCTGGGTACACCT	TACTTCAATTTTATTTGTTCCAAATTTCAAT	41941
Oy	1892	ATGCATATGACTGATGATTTGTCTCA	ATTTAGTGCAGAGGTGGGAAAAGGCTCGGA	1951
Db	41942	CTTATGTGCCGTATC-----TTA	CTTAATTAATTTGCAGATATACAGAACGAAT	41995
Oy	1952	TTGTGTATTTGCTGACGAGGTATAC	GGCAAGCTGTTCTTGGGACGCGCCGTTCA	2011
Db	41996	TTAATCTATTGCTGATGAAATATAT	ATGTCACCTTGTTATATGAAAGCACTCCT	42055
Oy	2012	ATGGGAATGTTTGGGCAATCACCCCT	GTGCTGTCCATATGAGGCTCTGTCCAAATC	2071
Db	42056	ATGGGTGTGTTTGGAGAGACTGTTC	CAAGTCACTCAATTGGAGACTATATCAAA	42115
Oy	2072	ATATGTGCTGGAATGCGGCTTGGAT	ATGGGTATGCGAGTGTACGACCCCGAAAG	2131
Db	42116	GTGTCTCCCTGGCTGGAATTTGGAT	GTATAGCAATTTGTATCTCTAAAGCAAT	42175
Oy	2132	GAATCTAAGGTAATT 2146		
Db	42176	GAATCAAGGTGATT 42190		

RESULT 12	AP003542/c	LOCUS	DEFINITION
AP003542	147640 bp	DNA	linear
Oryza sativa (japonica cultivar-group)	chromosome 6	clone F0589C03,	
*** SEQUENCING IN PROGRESS ***.			

ACCESSION	AP003542
VERSION	AP003542.1
KEYWORDS	GI:13810548
SOURCE	HTG; HTGS PHASE2.
	<i>Oryza sativa</i> (japonica cultivar-group)

ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehmerioideae; Oryzaceae; Oryza.

REFERENCE	AUTHORS	TITLE
1	Sasaki, T., Matsumoto, T. and Yamamoto, K.	<i>Oryza sativa</i> nipponbare (GA3) genomic DNA, chromosome 6, PAC

JOURNAL
Reference
2
147640
P0589C03
Published Only in Database (2001)
(bases 1 to 147640)

AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2001) Takuji Sasaki, National Institute of

Agrobiological Resources, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abr.affrc.go.jp, URL: <http://rpg.dna.affrc.go.jp/>).

COMMENT
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed

* NOTE: This is a 'working draft' sequence.

* by the finished sequence as soon as it is available and

* the accession number will be preserved.	
FEATURES	Location/Qualifiers
SOURCE	1. 147640

```

/organism="Oryza sativa" [ japonica cultivar-group ]
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="6"
/clone="P0589C03"

```

ORIGIN

Query Match	1.9%;	Score 213;	DB 2;	Length 147640;
Best Local Similarity	60.4%;	Pred. No. 1.3e-17;		
Matches 408;	Conservative 0;	Mismatches 255;	Indels 12;	Gaps 3.

1475 AAACAAGCGCGGTGGCAGAGCACTGTTCGAGGGCGTGCCTACATGCTATCGGCCGC 1534

Db 20788 AATTCAGAGCTATTGCAGAGTACCTATCTGTATCTTCTTACAGAAGCTTTGCCACAGAT 20777

1535 GACGCTTCCCTCACCCGCGCGGGACCAAGGCAATCGAGGTCAATAATCCCGGTGCTGGCC 1539

Db 20728 GATATTTTCCACATCTGGAGGTACCAAGCATCGAGTTGTTATGTCTGTTTTCGC 2066

1595 CAGACCGCCGGCCAACTTGTCTCCCAAGCCAGGCTAACCAACTACGAGGCGGC 1654

Db 20668 CA---ACAGGTGCCAATATATTGCTTCCAAAGCCGGGTACCCAAAACATGAAGCACAT 2061

1655 GCCCGCTTCAACAGGCTGAGGCTCCGGCATTTCCACCTCATCCCCCAACAGGGGTGGAG 1714

Db 20611 GCGGTCGTCACAGGATGGAGTGGCGCTCTATGATCTTGTTCACAGAGAGAGATGGGAG 20555

1715 ATGCACATCGACTCGCTGGAAATCCATCGCGCAGCAAGAACCCACCGCATGTCATCATTA 1774

D6 20551 ATTATGTTGAAGCTGTTGAACTTTACAGATGAAATCTGTTGCAATAGTACT 20499

07 1115 AACCCCAACACCCCGGCGGCGGCGGTACTCCACGACCATCTGCTCAAGT---TTC 1831

DB 20491 AACCCCAATACCTTCGTAAGTAAAGCATTATAGACATCTGCTCCAAAGGTGGTAGG 20493

[illegible]

20731 AAGGCGATCTGAATTGGTCAATTAGGTCCGCGGCGGCGAAAAAGCGTCGGAATA 1951

20371 CTATGTCGGCATC-----TTAACTTAGATTGAGATACAGCAAGCAAGTTCCGTCGTG 20311

1952 TTGTTGATTGCTGACGAGTATACGGCAAGCTGTTCTGGGACGCCCCCGTTTCATCCCA 2011

Db 20317 TTAGTCATTGCTGATGTAAGTAAATGGTCACTTTGTTATGGAAGCACTCCTTTGTGCA 20355

2012 ATGGGAGTGTTTGGGCATCACCCTGCTGCTCCATAGGCTCTGTCCAAGTCATGG 2071

Db 20257 ATGGGTGTGTTGGAGACTGTTCCAGTACTCAATTGGGAGCTATATCAAAGAGATGG 2019

2072 ATAGTCCTGATGCGGCTTGATGGGTAGCGGTGTACCACCCCAGAAAGATCTTACAG 2131

Db 20197 GTTGTCCTGGCTGAGATTGGATGATAGCAATTGTGATCTCTAAGGCAATTCTGAAA 20133

2132 GAACTAGGTACTT 2146
|||||

Db 20137 GAACAAAGGTGATT 20123

RESISTIT 13

	ATF2009	LOCUS	ATF2009	DNA	linear	PLN 03-AUG-1999
	93695 bp					

DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone F2009 (ESSA project).

ACCESSION	1	GI:2842474
VERSION	2	
	AL021749	
	AL021749.1	

KEYWORDS
SOURCE
Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
REFERENCE Bevan, M., Rose, M., Hempel, S., Entian, K.-D., Hohleisel, J.,
Mewes, H.-W., Meyer, K. F. X., Lemcke, K. and Schueller, C.
AUTHORS Unpublished
2 (bases 1 to 93695)
JOURNAL EU Arabidopsis sequencing project.
REFERENCE Direct Submission
AUTHORS Submitted (30-JUN-1999) MIPS, at the Max-Planck-Institut fuer
TITLE Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
JOURNAL schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.
FEATURES
source
1. 93695
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
18. . 467
/gene="P2009.10"
complement (18. . 467)
/gene="P2009.10"
complement (18. . 467)
/gene="P2009.10"
/codon_start=1
/product="hypoetical protein"
/protein_id="CA16872.1"
/db_xref="GI:2842475"
/db_xref="UniProt/TREMBL:O49442"
/translation="MTSYKLPKRVISKRRTACFLVYVAILSSNMSINSTVWY
ENHQTSSGLPAIYASVLGAVFGVLSMAALPIAIPAVIIVITISVVTMAIRGSRKR
VVIKRVTKKEIGFVFLVLEKGNFVLLCALLVFFVFNSSSS"
complement (18. . 467)
/gene="P2009.10"
/number=1
2196. . 2675
/gene="P2009.20"
complement (2196. . 2675)
/gene="P2009.20"
complement (2196. . 2675)
/gene="P2009.20"
/codon_start=1
/product="hypoetical protein"
/protein_id="CA16873.2"
/db_xref="GI:5701784"
/db_xref="UniProt/TREMBL:O49443"
/translation="MADKEBOTMSTYFLRVISKRRTWVCLPLVVVAVILSSRNS
LNSITVWGENHQTSSGLPAIYASVLGAVFGVLSMAALPIAIPAVIIVITISVVT
APVYSRKRVIVISKRRTKEIGVIFKVLKGNFVLLCAVIAIVFFVFNSSSS"
complement (2196. . 2675)
/gene="P2009.20"
/number=1
4143. . 6188
/gene="P2009.40"
join(4143. . 4961, 4987. . 5044, 5096. . 6188)
/note="similarity to lectin receptor-like serine/threonine
kinase IecR1, Arabidopsis thaliana, PIR2:56589
Contains Protein Kinases signatures and profile, pos.
328-349 & pos. 444-456"
/codon_start=1
/product="receptor protein kinase like protein"
/protein_id="CA16875.1"
/db_xref="GI:2842478"

/db_xref="GOA:O49445"
/db_xref="UniProt/TREMBL:O49445"
/translation="MFSKVSITLPSLASLILFRSTTGIEFTYNSNFTTTTLGLANAT
VKSPPSITLTNQTTPTISGRLYPBRINASSSSAPLPEAFSFLPSMAPFRLSGHG
FAFVFLPSESESAASSQHLGFNFTNNGDNNRSIFAVEPFAVNOENDINDNVGV
DVSILTSVASETAGVGGDQRFETELKNGENQAVIENFGSAINTVMAASARKP
IRPLISIPINTGLVLDMPFVGTASGOLVQSHRILMSNSNSFSGDVSLSKGF
IAGVSGVVLQRLGSDVDEMETEYVPHRYVKDYLAEKGFSDENMIGVGNKRYRG
VLEKRAVAKRIMNSPRESVATSEFLAVSLSGLRLRKNIYGLKMSKKGESITLI
YEYENSGVDNRIFDCEMLNWEERMYIRDLASGMLYLHGEWETKVLHRIKSSNVL
LDKDMNARVGDFFGLAKLQNTSEKAVSTTHVGTAGYMAPELVKTGRSAQTDVYSPG
FLVEVCCRRPIEERREGIVEMIGLMEKDKVVDGLDERIKANGVFAVEEYEMALRIG
LFLVCHDPRVAPKRGVOVQVLEOGRLYVDGERETSLERVKSSTLETGTGSRQGH
TFQVWNSSVSNSEFQYDILHHR"
4143. . 4961
/gene="P2009.40"
/number=1
4942. . 4986
/gene="P2009.40"
/number=1
4987. . 5044
/gene="P2009.40"
/number=2
5045. . 5095
/gene="P2009.40"
/number=2
5096. . 6188
/gene="P2009.40"
/number=3
6559. . 11254
/gene="P2009.30"
complement (join(6559. . 6669, 6878. . 7050, 7305. . 7428,
7537. . 7647, 7763. . 7915, 10629. . 10733, 10823. . 10875,
11074. . 11254))
/gene="P2009.30"
complement (join(6559. . 6669, 6878. . 7050, 7305. . 7428,
7537. . 7647, 7763. . 7915, 10629. . 10733, 10823. . 10875,
11074. . 11254))
/gene="P2009.30"
/note="similarity to ENOD20 gene (small copper-binding
protein), Medicago truncatula, PATCH: E258656"
/codon_start=1
/product="copper-binding protein-like"
/protein_id="CA16874.2"
/db_xref="GI:5701785"
/db_xref="GOA:O4944"
/db_xref="UniProt/TREMBL:O49444"
/translation="MGLVMPEDLYIMFVLMGLGFTISNGYKPYVGKDGCVPTPSBD
YSHSHNRFCVNDTLPLQVPTFSLVATLKLVEISSPGSPGVSESDNSRVAP
GPATAPNSAGIGISTROLOASBPVSSPALLSGKEBKXITPKROKVAATK
SIKSPKKNVIALVYRGMRVEDALLIQLVYKRAAQTVYRHTTCFVNTMLKGYKTI
LQVTHAANATHNHGIDPDKLVIVIRETTPSEAEIARLKVNFKKSKREKQVLP
HKLIETSPINRRGTKANHRSELVLTIIIDVTCVGNMKKRLDNLTNQNNIYHNHPS
GP"
complement (6559. . 6669)
/gene="P2009.30"
/number=1
complement (6670. . 6877)
/gene="P2009.30"
/number=1
complement (6878. . 7050)
/gene="P2009.30"
/number=2
complement (7051. . 7304)
/gene="P2009.30"
/number=2
complement (7305. . 7428)
/gene="P2009.30"
/number=3
complement (7429. . 7536)
/gene="P2009.30"
/number=3
complement (7537. . 7647)


```

      /gene="F2009.30"
      /number=4
      complement(17648..7762)
      /gene="F2009.30"
      /number=4
      complement(17763..7915)
      /gene="F2009.30"
      /number=5
      complement(17916..10628)
      /gene="F2009.30"
      /number=5
      /number=5
      /gene="F2009.30"
      /product="rRNA-Lys"
      /note="codon recognized: AAA"
      complement(10629..10733)
      /gene="F2009.30"
      /number=6
      complement(10734..10822)
      /gene="F2009.30"
      /number=6
      complement(10823..10875)
      /gene="F2009.30"
      /number=7
      complement(10876..11073)
      /gene="F2009.30"
      /number=7
      complement(11074..11254)
      /gene="F2009.30"
      /number=8
      /note="ESTmatch (T04156_F)"
      12220..12639
      /gene="F2009.50"
      join(13176..13187,13379..13519,13534..13707,14167..14233,
      14308..14385,14487..14635,14753..14783,15222..15422,
      15885..16117,16178..16213,16388..16586,16629..16739,
      16804..16811)
      /gene="F2009.50"
      /codon_start=1
      /product="hypothetical protein"
      /protein_id="CA16876.2"
      /db_xref="GI:5701786"
      /translation="MESKILFGRKQBAEVPSPAMNITGYRGPHALLHFEVYFYS
      GLSTMNTNVDYLSNTRVRSFGKTPKINSKRPDPKXGNSVVELVAIPTFI
      TGVHYVQIGTFSQVQESPRDLKRLKLSFVNEKXCHIEIQAQVSRVAASENMD
      1.9%; Score 212.6; DB 8; Length 93695;
      Beat Local Similarity 53.2%; Pred. No. 1.4e-17;
      Matches 577; Conservative 0; Mismatches 459; Indels 49; Gaps 4;

```

```

      Db 33617 TACTTGTTCACCCCACTGTAAGTCTGTAAACATTCAGAAACCTAAGATATTTTCATGCA 33676
      Qy 7013 TTATATAAAACAAACACGCCCTGACAGACACTTGTACAGAGCCGCTTACAAGCTAT 7072
      Db 33677 ATGTGTTGTTAAACAGAGCGGCTGCAAGATTACCTGAACCGAAGCTTGAGAACAGTA 33736
      Qy 7073 CGCCGACGACGCTCTCTCAACCGCGGCGAAGCTGAGCGATGAAAGTCATTAATCCCG 7132
      Db 33737 AGCTTAATATATGATTAATCATTAAGTGTGGAATGCAACCAAGGAAAGTAAGTAGTCTTCA 33796
      Qy 7133 TCGTGGCCCAAGACTGCGGCGCCAACTACTGTGTTCCCGGCGAGCTATCCAAATTAAG 7192
      Db 33797 CACTGGCTC---GACCAACGCTTAACATCTCTTCCAAAGGCTTAATACCTCCTACATAG 33853
      Qy 7193 AGCGGAGAGCGGACATTCACAAAGCTGAGGTCCGCACTTCGACTATCCCGCAAGG 7252
      Db 33854 AGGCTGTGCGGCTGCTACAGTGAAGCTGAGGTCCGCAAGTTCGATCTCTCCAGAGAAAG 33913
      Qy 7253 GGTGGAGATGCAATGCAATGCACTGCTGGAATTCATGCCGACAGAAACACCAACCGCATGG 7312
      Db 33914 AATGGAGATGATGATCTCCAGAGCATGAAGCCATGGAATGAGAACAGATTGCATGG 33973
      Qy 7313 TCATCATTAACCAACCAATCCGTGCGGACAGCTTTACTCTTACGACATCTGCGCAAG 7372
      Db 33974 TTATCATTAACCTTAACCAACCCCTGTGGAATGTTTACTCTTAAGATCATCTTAAGAAAG 34033
      Qy 7373 TTTTGATTCATGATCATCTCTGCTGTGATGATGACCGGCTGTGGAATGATATGATG 7432
      Db 34034 TCCGTATTC-----ATACCTTTTAATTAAGCTTTGACTCGTGAAGACTGAAA 34080
      Qy 7433 GATTGCGTTGCTAATGCTGTGCTGATGATGCTGTTGTTATCAGTCCGAGAGGTGGC 7492
      Db 34081 TATTGGTGTGATGA-----ATGAACCTTTAATGAGTGGCGAGACGGC 34124
      Qy 7493 AAGGAAGCTGGAAATTTGTTGATTCGCTGACAGAGTTTACGCAAACTGTTCTGGGACG 7552
      Db 34125 TAAGAAAGCTGGAAATGATGATATCAAGACGAAGTTTATGCAAAACATCTTCGAGA 34184
      Qy 7553 CGCCCGTTTATCCGATGAGGCGCTTTGGGCAATTGCCCGGCTGTGCAATGGATC 7612
      Db 34185 CAATCTTTGTCCAAATGGGAGATTTTCTTGATATATCCGATTATCACTTTGGGCGG 34244
      Qy 7613 TCTGTCAAGTCTGTGATGATGCTGATGATGACGACTTGAATGGTGGCGGTGACAGCC 7672
      Db 34245 TATATGAAAGATGATGATTTCTGCTGTGGAATTTGGCTGATGCTTGAATGACCC 34304
      Qy 7673 CACAAAGATTTAGAGAAACTAAGTAGCTTTAGCTCCCTATCATTTCTCATATGCT 7732
      Db 34305 CAGAGGAATCTCAAGTCCAGAGGCTGTGCTCTTTTACTGATCATCATTAAC 34364
      Qy 7733 ACTGT 7737
      Db 34365 TGTGT 34369

```

RESULT 14
 ATCHR168
 LOCUS
 DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 68.
 ACCESSION AL161572
 VERSION AL161572.2 GI:7269643
 KEYWORDS
 SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (bases 1 to 5241; 88721 to 138031)
 Lemnard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
 Mewes H.W., Lemcke K. and Mayer K.F.X.
 JOURNAL Unpublished
 REFERENCE 2 (bases 5137 to 88829)

AUTHORS Pohl, T., Weizenegger, T., Mewes, H. W., Lemcke, K. and Mayer, K. F. X.
JOURNAL Unpublished
REFERENCE 3 (bases 137918 to 199749)
AUTHORS Rose, M., Hempel, S., Entian, K.-D., Mewes, H. W., Lemcke, K. and Mayer, K. F. X.
JOURNAL Unpublished
REFERENCE 4 (bases 1 to 199749)
AUTHORS EU Arabidopsis sequencing project.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemccke@mips.biochem.mpg.de, mayermips.biochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7JY Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/chal/> this fragment has an overlap with ATCHR167 at the 5' end and an overlap with ATCHR169 at the 3' end.
FEATURES
source location/Qualifiers
1. 199749
/organism="Arabidopsis thaliana"
/mol_type="Genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
6341..8260
/gene="AT4g27900"
join(6341..6448,7253..7708,7926..8066,8180..6260)
/gene="AT4g27900"
/note="similarity to zinc-finger protein S12569, Oryza sativa, PIR3:JE0115
containing EST gb:142849, T43671, A1995239.1, T04160, R30019, T22597, T22233"
/codon_start=1
/product="putative protein"
/protein_id="CAB79593.1"
/db_xref="GI:7269644"
/db_xref="UniProt/TREMBL:Q9SUE8"
/translation="MYAETSLAFRPMQSSSPETIHQFEDLFKSYKLSDEMNLYEASEYDFGEISDLFPAPEPIIEPMLAVDPLSQELVGLDGLSQDQILDKAFYDEQDLVKSAMSEPLSEVLDIKNISLVAKNDNVKTSYVSDVPIPKSYSGNLSMNVESHKDVAITGSPDPDPVDYGMWRAPASSEDITGLTGNTGNYQSOLDRIITISCTSEDDREKLSRYRNKSKRPFGRKIKYACRKALDSQPRIRGRPAKTEWQK"
6341..6448
/gene="AT4g27900"
/number=1
6449..7252
/gene="AT4g27900"
/number=1
7253..7708
/gene="AT4g27900"
/number=2
7709..7925
/gene="AT4g27900"
/number=2
7926..8066
/gene="AT4g27900"
/number=3
8067..8179
/gene="AT4g27900"
/number=3
8180..8260
/gene="AT4g27900"
/number=4
10177..15419
/gene="AT4g27910"
join(10177..10956,11042..11097,11166..11403,11496..11624,11885..11953,12055..12314,12474..12610,12836..13025,13575..13720,13825..13929,14174..14238,14331..14541,14835..14897,14950..15039,15129..15215,15303..15419)
/gene="AT4g27910"

/note="similarity to MLT protein, Gallus gallus, Av011003
containing EST gb:217631"
/codon_start=1
/product="putative protein"
/protein_id="CAB79593.1"
/db_xref="GI:7269645"
/db_xref="GOA:Q9SUE7"
/db_xref="UniProt/TREMBL:Q9SUE7"
/translation="MYIKRKKTQIPISLERKLNESKRRKRNLTGGGGYYPNLT
GEIAGIVPGNGRNGFSASCTEYTKPVEVESLSKRSDSGITVDSPPAVSRPLV
RTSGRQVLPESRNDVSLDWRVDSKDCDLEEEIECRNEKXVSPFVPAWNLKSK
ELDKRSYALCKEERFHECHNDARARVDEKLPNKGTPEPENSGLVMAKGRN
EPPFPAIYIDMTAPELVLSGCI PDAAVYFPGHSGNENRDYAMVRGAIPEVDY
VAROQPEIQGCKPGRNOMLREAPLADQFTEKLMDHILAAIGNSTPDSFPMIO
ETAVSNGELNNAPRQASVSNITLILKKRPLACAGETVISEMAKKMDLPGDQ
LICKPCSRVRCDCGKVRINHACDQISPHLKITVLAEOVLTLYQIQKRTSQSR
PLRNIFYCYFLNGELLAKCLNTGVFWRYAKGQVLPKVVIVCGVGVFEPRL
HLVVCCKGSCGPKKKALSEMERHSGSKSKMTSVKYSKLALIEDWKEKESCNML
QCPCFCDAFYGFSPFCQMRMSRMSRSGQNDVQPDGLKITTRLISATGALKP
TDVETLWVHTYCAPOPEVCPASREKKEPAGIISPTNPKVCVICKOIHGSCQTC
CKCSTYTHAMCASRAGTMEHLCERKNGQITTVNVSICATYRARPNPVLITQPSGA
FSASLVQNKKKGSRILSLIRDEDAPEVNTTCDPESARCVFPKKLNSKGTIS
PLAKNSLIMEFRDNNVQRTEDRVCFGRSGIHGGLPARNMIQGEVWLEVRGQVRG
SIADLRARYRVRVXKDCYLFKISEVVVDATDKNIRALINHVSCLAMTYVQEI"
10177..10956
/gene="AT4g27910"
/number=1
10957..11041
/gene="AT4g27910"
/number=1
11042..11097
/gene="AT4g27910"
/number=2
11098..11168
/gene="AT4g27910"
/number=2
11169..11403
/gene="AT4g27910"
/number=3
11404..11495
/gene="AT4g27910"
/number=3
11496..11624
/gene="AT4g27910"
/number=4
11625..11884
/gene="AT4g27910"
/number=4
11885..11953
/gene="AT4g27910"
/number=5
11954..12054
/gene="AT4g27910"
/number=5
12055..12314
/gene="AT4g27910"
/number=6
12315..12474
/gene="AT4g27910"
/number=6
12475..12610
/gene="AT4g27910"
/number=7
12611..12835
/gene="AT4g27910"
/number=7
12836..13025
/gene="AT4g27910"
/number=8
13026..13574
/gene="AT4g27910"
/number=8
13575..13720

intron	/gene="ATG27910" 13721. .13824 /gene="ATG27910" /number=9
exon	13825. .13929 /gene="ATG27910" /number=10
intron	13930. .14173 /gene="ATG27910" /number=10
exon	14174. .14238 /gene="ATG27910" /number=11
intron	14239. .14330 /gene="ATG27910" /number=11
exon	14331. .14541 /gene="ATG27910" /number=12
intron	14542. .14834 /gene="ATG27910" /number=12
exon	14835. .14897 /gene="ATG27910" /number=13
intron	14898. .14949 /gene="ATG27910" /number=13
exon	14950. .15039 /gene="ATG27910" /number=14
intron	15040. .15128 /gene="ATG27910" /number=14
exon	15129. .15215 /gene="ATG27910" /number=15
intron	15216. .15302

Query Match	Similarity	1.9%	Score	212.6	DB	8	length	199749
Best Local	Similarity	53.2%	Pred. No.	1.4e-17				
Matches	577	Conservative	0	Mismatches	459	Indels	49	Gaps
								4
Qy	6653	CGGGCGAAGACGCGCATCTCTGGCGACGACGCGGGCGAAGAACGATCCGGCGATACGG	6712					
Db	171251	CGTTTCAGGGGAAAGCACAACGACGACCAAGACCTCCAGCGTCAAGATGAGATCATCGTC	171310					
Qy	6713	TACAGATATAGGCGGACGCTGAGAGAGAGCGGGCGCGCGCGCGTCCGCGCGCTGCGCCAC	6772					
Db	171311	TACAAAGCTTTGTGATGATATGCAAGCTTGATGTGAAAAAGCTCTTTTATCCCTTGCTCAC	171370					
Qy	6773	GGTGACCCGTCCTGTTCCCGGCTTCGCGACGCGCGCTGAGGCGGAGACGCGGTAC	6832					
Db	171371	GGTGACCTTTCTGTCTACCTTTGTTACCGCACTTCATCTCGTCGAGAAACGCGGTGTC	171430					
Qy	6833	GCGCGCTGCGGACGCGGCGCATTTCACTGTAGCGCGCGCGCTCGGCTCTCCCGCGCA	6892					
Db	171431	GACGTCTTCGTGCTCGGCAAGGGTAACTTTAGCGCCCGCGCGGAATTTCTCCAGCC	171490					
Qy	6893	CGAAGTAACTTTACAGCTTCAACCGTAATGTATGCTGAGCATGATGCGCGGTTTAC	6952					
Db	171491	AGACAGTAAATATACATAATTTTTCTCTTACTTC-----TTAAT	171533					
Qy	6953	TTAGTGCCTCGCGCTGTTCTTCCCGGTGCGTTCAAAATTTTAACCTTCTATAGTACC	7012					
Db	171534	TACTTTGTTCCACCACTGTAAAGTCTGTAAACATTCAAAACCTAAACGATATTTTCATGGA	171593					
Qy	7013	TTATAAACAACAACGCGCGGTAGACAGACACTGTCAACAGGCGTCCCTAACAGTAT	7072					
Db	171594	ATGTTTTTTGTTATCAAGGCGGTGCGAGATTTAGTGAAACGAAGCTTGACGAACAAGTAA	171653					
Qy	7073	CGGCGACGACGTCTTCTCACCGCGGCGGAATCAGGCGATGGAATCATATCCCG	7132					

Db	171654	AGCCTAATGATGTATTCATTA	CTGTCGGATGCAACCAAGGATAGAAAGTATGTTCTTCAAGT	171713
Qy	7133	TGCTGGCCCAAGCTGCGGCGCA	CAATACTGCTTCCCGGCGAGGCTATCCAAATTACG	7192
Db	171714	CACGTGGCTC---	GACCAAAAGCTAACATCTCTTCTCCAAAGGCGTGGTACCTCACTACG	171770
Qy	7193	AGGCGCGAGCGGCATTTCA	CAACAAGCTGGAAGTCCCGGCACTTCGACCTCAATCCCGACAAAG	7252
Db	171771	AGGCTGCGCGGTCTTACAGTGA	CTCGAGGTCCGAAGTTCATCTCTTCCAGAAAG	171830
Qy	7253	GGTGGGAATCGACATCGACTCG	CTGGAATCCATCGCGACAAAGAACACCGCGATGG	7312
Db	171831	AATGGGAATTTGATCTCCACG	GCATCGAAGCCATGGCAGATGAGAACACAGTTTGCATGG	171890
Qy	7313	TCATCATAAACCCAAACAAT	TCGTGCGGACGCTTACTCTTACGACCATCTGGCCAAAG	7372
Db	171891	TTATCATTAACCTCTAACAA	ACCCCTGTGAAATGTTTACTCTTACGATCATCTTAAGAAAG	171950
Qy	7373	TTTTCGATCCATGCATCCT	CTGCTCGTTGATCGACCGGTCTGTTGAACATATGATAG	7432
Db	171951	TCGCTATTC-----	ATACTTTATATGCTTTGACTCGTGAAGACTCGAAA	171997
Qy	7433	GATTCGCTTGCTATCGTGTG	TCTGATGCTGTTGGTATTCAGTGCAGGAGTGGC	7492
Db	171998	TATTGTGCTATGA-----	ATGAACCTTTATATAGTGGCGGAAGCGC	172041
Qy	7493	AAGGAAGCTCGGAATATTTG	TGTGATCGCTGACGAGGTTTACGSCAAACGTGTTCTGGGCAAG	7552
Db	172042	TAAAGAACTGGGAATATATG	TATATCAACAGCAAGATTTTATGCCAAACAATCTTCGGAGA	172101
Qy	7553	CGCCCGGTTATCCCGATGGG	CGTCTTGGGCAATGCGCCGCGTGTGCATTTGGATGC	7612
Db	172102	CAAACTTTTGTCCCAATGGG	GAAGTTTCTTGATTAATCCGGTTATCACTTTGGCGGG	172161
Qy	7613	TCTGTCCAAGTCGTGATAGT	GCCTGATGCGCACTTGGATGGGTGGCGGTATACGACCC	7672
Db	172162	TATATCGAAAGATGATTTGT	TGTTCTCGTGTGGAGATTTGGCTGGATCGCTTTGAATGACCC	172221
Qy	7673	CACAAAGATTTTAGAGAAA	CTAAGTAAAGCTTAACTCCCTATCATTTCTTCAATATGCT	7732
Db	172222	CAGAGAAATTCACAAGTCC	ACAGAGGGGTGTCTTCTCTTTTATAGTATCATCATTAACCT	172281
Qy	7733	ACTGT 7737		
Db	172282	TGTGT 172286		

RESULT 15				
LOCUS	AC150698/c			
DEFINITION	Oryza sativa (japonica cultivar-group)	DNA	linear	PLN 29-SEP-2004
ACCESSION	OSUNBA0072P10,	complete sequence.		
VERSION	AC150698			
KEYWORDS	AC150698.3	GI:52839695		
SOURCE	HTG.			
ORGANISM	Oryza sativa (japonica cultivar-group)			
REFERENCE	Oryza sativa (japonica cultivar-group)			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
	Ehretidoideae; Oryzaceae; Oryza.			
	1 (bases 1 to 147540)			
	Dalal,V., Pandit,A., Ghazi,I.A., Yadav,M., Singh,P.K., Singh,A.,			
	Mendiratta,S.D., Singh,H., Srivastava,S.K., Singhal,V., Dixit,R.,			
	Batra,K., Galikwad,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.			
TITLE	Oryza sativa (japonica cultivar-group) chromosome 11 BAC clone			
JOURNAL	OSUNBA0072P10			
REFERENCES	Unpublished			
AUTHORS	2 (bases 1 to 147540)			
	Yadav M., Pandit A., Galikwad K., Srivastava S.K., Singh H.,			
	Singh A., Mendiratta S.D., Singh P.K., Singhal V., Dalal V.,			
	Pai, A.K., Ghazi I.A., Dixit R., Batra K., Sharma T.R., Mohapatra T.			
	and Singh N.K.			

TITLE Direct Submission
JOURNAL Submitted (01-AUG-2004) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India
REFERENCE 3 (bases 1 to 147540)
AUTHORS Yadav,M., Pandit,A., Gaikwad,K., Srivastava,S.K., Singh,H., Singh,A., Mendiratta,S.D., Singh,P.K., Singhal,V., Dalal,V., Pat,A.K., Ghazi,I.A., Dixit,R., Batra,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2004) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India
REFERENCE 4 (bases 1 to 147540)
AUTHORS Dalal,V., Pandit,A., Ghazi,I.A., Yadav,M., Singh,P.K., Singh,A., Mendiratta,S.D., Singh,H., Srivastava,S.K., Singhal,V., Dixit,R., Batra,K., Gaikwad,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2004) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India
COMMENT On Sep 29, 2004 this sequence version replaced gi:50897315.
FEATURES
Source Location/Qualifiers
1..147540
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="11"
/clone="OSUNBA0072P10"
/clone_11b="HindIII"
ORIGIN
Query Match 1.9%; Score 209.6; DB 8; Length 147540;
Best Local Similarly 56.6%; Pred. No. 3.5e-17;
Matches 429; Conservative 0; Mismatches 324; Indels 5; Gaps 2;
QY 1457 ATACCAACAAACAAACAAACAAACACACGCGCGTGCAGACGACCTGTGCGAGGGGTGCGG 1516
DB 52441 ATCTTAACAACTGAAGTTGAATTCAGAGCTTATCCGCGGATCTTATGCGAGACTTGCCA 52582
QY 1517 TACATGCTATCGCGGACGAGCTTCTCTCAACCGCGCGGACCCAGCGCATGAGATC 1576
DB 52581 TATGAGCTATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 52522
QY 1577 ATATATCCCGGTGCTGCGCGACGCGCGCGCAACATTTGCTCCCAAGCCGACGCTAC 1616
DB 52521 ATCTGCTCTGCTGATGCTGCG---CCCTGCTGCAACATCTGCTGCCAAGCGCGGTAC 52465
QY 1637 CCAAACTACGAGCGCGCGCGCGCTTCAACAGGCTGAGAGTCCGGCATTTGACCTATC 1696
DB 52464 CTGTTCCAGAGGACGCGCGCGCTTCAATGAGTGAAGGTCAGTACTTTGATCTTCTC 52405
QY 1697 CCCGACAAAGGGGTGGAGATGACATGACCTGCTGGAATTCATCGCCGACCAAGAACCC 1756
DB 52404 CCAAGAGTGGCTGGAGGTTGATCTTGAAGAGTGCAGAACTTGCTGACCAACACG 52345
QY 1757 ACCGCTATGCTCATTAACCCCAACACCCGCGCGAGCGTTTACTCTTACGACAT 1816
DB 52344 GTTGCAATGCTCATTAACCCCAACACCCGCGCGAGTCTTCACTTCTGACAT 52285
QY 1817 CTGTCAAGGTTTCAATCTCTTCTGCTGGAATGATGATGATGATGATGATGATGATGATG 1876
DB 52284 TTGGCCAAAGCAAGTACTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 52225
QY 1877 GAATTTCTTTTTCGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1936
DB 52224 GGAATTTCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 52167
QY 1937 AAAAGGCTCGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1996
DB 52166 AAAAGGCTCGCATTTTGTTCATGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 52107

QY 1997 GCCCGTTCAATCCCAATGGAGGTGTTGGGACATCAACCCCTGCTGCTCATAGGCT 2056
DB 52106 AATTAATTTGTGCGGAGTGGGGGTGTTGGCTCACTAGCTCAAGTTCTTACCTTAGATCG 52047
QY 2057 CTGTCAAGTCAATGATAGTCCCTGATGCGGCTTGAATGGGTAGCGGTACGACCCC 2116
DB 52046 ATATCGAAGAGATGGGTGCTGATGCGGACTTGATGCGACTTGATGATGATGATGATGATG 51987
QY 2117 AGAAGATCTTACGAGAACTTACGATCTTAATCTATATCTTTTCAATGCTA 2176
DB 51986 AATGCGCTATTTCAAGACCAAGAGTGTATTAAGTTTAACTTCAACCCCTTTCATCAA 51927
QY 2177 CTAGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2214
DB 51926 GTATTTGCAATCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 51889

Search completed: October 15, 2005, 01:33:35
Job time : 44828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 13:04:36 ; Search time 98.0375 Seconds

Title: US-10-019-783-3 (without alignments)
Sequence: 1 tctggagatcagtgatcccaaca 24 11862.046 Million cell updates/sec

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hng:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sta:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9416466	100%	147540	US-10-019-783-3	Direct Submission

1	24	100.0	10966	6	E49893	E49893 Creation of
2	24	100.0	10966	6	BD012010	BD012010 Creation
3	24	100.0	10966	6	AB024006	AB024006 Hordeum v
4	19.4	80.8	2969	2	AC022300	AC022300 Mus muscu
5	19.4	80.8	104147	2	AC004703	AC004703 Homo sapi
6	19.4	80.8	152056	9	AC146691	AC146691 Macaca mu
7	19.4	80.8	171952	10	AC002324	AC002324 Mus muscu
8	19.4	80.8	172940	10	AC126544	AC126544 Homo sapi
9	19.4	80.8	192550	2	AC025584	AC025584 Mus muscu
10	19.4	80.8	193004	10	AL591177	AL591177 Mouse DNA
11	19.4	80.8	198388	2	AC073426	AC073426 Homo sapi
12	19.4	80.8	206578	2	AC126546	AC126546 Homo sapi
13	19.2	80.0	57662	2	AC107969	AC107969 Homo sapi
14	19.2	80.0	85378	2	AL356577	AL356577 Homo sapi
15	19.2	80.0	98360	9	HSDJ247C2	HSDJ247C2 Human DNA
16	19.2	80.0	135353	9	AC146086	AC146086 Pan trogl
17	19.2	80.0	145264	9	AC107939	AC107939 Homo sapi
18	19.2	80.0	154344	9	AC073308	AC073308 Homo sapi
19	19.2	80.0	168139	2	AC016936	AC016936 Homo sapi
20	19.2	80.0	185259	2	AC063953	AC063953 Homo sapi
21	19.2	80.0	187204	9	AC084082	AC084082 Homo sapi
22	19.2	80.0	188417	2	AC150805	AC150805 Callithrix
23	19.2	80.0	219730	2	AC110957	AC110957 Rattus no
24	19.2	80.0	235033	2	AC125338	AC125338 Rattus no
25	19.2	80.0	235082	2	AC094542	AC094542 Rattus no
26	18.8	78.3	150	6	AX753668	AX753668 Sequence
27	18.8	78.3	500	6	AX753680	AX753680 Sequence
28	18.8	78.3	1026	6	CO745253	CO745253 Sequence
29	18.8	78.3	1689	6	AX753658	AX753658 Sequence
30	18.8	78.3	1826	6	CO719286	CO719286 Sequence
31	18.8	78.3	1867	6	AX082301	AX082301 Sequence
32	18.8	78.3	1932	6	BC014126	BC014126 Homo sapi
33	18.8	78.3	2271	6	AX576068	AX576068 Sequence
34	18.8	78.3	2610	6	AX753655	AX753655 Sequence
35	18.8	78.3	2830	6	BD270045	BD270045 Sequence
36	18.8	78.3	3173	6	AX753654	AX753654 Sequence
37	18.8	78.3	3248	6	AX753657	AX753657 Sequence
38	18.8	78.3	3418	9	BC037539	BC037539 Homo sapi
39	18.8	78.3	5296	9	AK074084	AK074084 Homo sapi
40	18.8	78.3	91200	9	AP001152	AP001152 Homo sapi
41	18.8	78.3	147556	2	AC011007	AC011007 Homo sapi
42	18.8	78.3	154324	2	AC122552	AC122552 Mus muscu
43	18.8	78.3	173235	2	AC023756	AC023756 Homo sapi
44	18.8	78.3	204954	10	AC116151	AC116151 Mus muscu
45	18.4	76.7	174965	9	AC036222	AC036222 Homo sapi

ALIGNMENTS

RESULT 1	E49893	10966 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	E49893	Creation of iron-deficiency	resisting rice plant.		
DEFINITION	E49893	Creation of iron-deficiency	resisting rice plant.		
ACCESSION	E49893.1	GI:18629372			
VERSION	JP 2001017012-A/1.				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 10966)				
AUTHORS	Mori,S., Nakanishi,N., Takahashi,M. and Nishizawa,N.				
TITLE	Creation of iron-deficiency	resisting rice plant			
JOURNAL	Patent: JP 2001017012-A 1 23-JAN-2001;				
COMMENT	SCIENCE & TECH AGENCY,SATOSHI MORI				
	OS Hordeum vulgare L. var. Igr1				
	PN JP 2001017012-A/1				
	PD 23-JAN-2001				
	PR 05-JUL-1999 JP 1999190318				
	PI SATOSHI MORI,NOBUHITO NAKANISHI,MICHIKO TAKAHASHI, PI				
	NISHIZAWA				
	PC A01H5/00,C12N5/10,C12N15/09//C12N5/10,C12N5/00,C12N15/00 CC				

FEATURES	source	1..10966	Location/Qualifiers
			/organism="Hordeum vulgare L. var. Igr1".
ORIGIN			
Query Match	100.0%; Score 24; DB 6; Length 10966;		
Best Local Similarity	100.0%; Pred. No. 0.055;		
Matches	24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 TCTGGATCTCAGTCATCCACA 24		
Db	8475 TCTGGATCTCAGTCATCCACA 8498		

RESULT 2	BD012010	10966 bp	DNA	linear	PAT 02-AUG-2002
LOCUS	BD012010	Creation of iron-deficiency resisting rice plant.			
DEFINITION	BD012010	Creation of iron-deficiency resisting rice plant.			
ACCESSION	BD012010.1	GI:22092199			
VERSION	WO 0101762-A/1.				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 10966)				
AUTHORS	Mori,S., Nakanishi,H., Takahashi,M. and Nishizawa,N.				
TITLE	Creation of iron-deficiency resisting rice plant				
JOURNAL	Patent: WO 0101762-A 1 11-JAN-2001; JAPAN SCIENCE AND TECHNOLOGY CORP., SATOSHI MORI, HIROMI NAKANISHI, MICHIKO TAKAHASHI, NAOKO NISHIZAWA				
COMMENT	OS Hordeum vulgare L. var. Igr1 PN WO 0101762-A/1 PD 11-JAN-2001 PF 04-JUL-2000 WO 2000JP004425 PR 05-JUL-1999 JP 99P 190318 PI SATOSHI MORI, HIROMI NAKANISHI, MICHIKO TAKAHASHI, PI NAOKO NISHIZAWA PC A01H5/00, C12N5/14, C12N15/52 CC				
FEATURES	source	1..10966 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"			
ORIGIN		Location/Qualifiers.			
		FH Key Location/Qualifiers.			
Query Match	100.0%; Score 24; DB 6; Length 10966;				
Best Local Similarity	100.0%; Pred. No. 0.055;				
Matches	24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 TCTGGATCTCAGTCATCCACA 24				
Db	8475 TCTGGATCTCAGTCATCCACA 8498				
RESULT 3	AB024006	10966 bp	DNA	linear	PLN 27-NOV-1999
LOCUS	AB024006	Hordeum vulgare naat-B and naat-A genes for nicotianamine			
DEFINITION	AB024006	amino transferase, complete cds.			
ACCESSION	AB024006.1	GI:6469088			
VERSION	AB024006.1	nicotianamine amino transferase.			
KEYWORDS	Hordeum vulgare subsp. vulgare				
SOURCE	Hordeum vulgare subsp. vulgare				
ORGANISM	Hordeum vulgare subsp. vulgare				
	Bkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Gymnamiales; Magnoliopsida; Liliopsida; Poales; Poaceae;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				


```

repeat_region /rpt_family="AT_rich"
25785..26070
/rpt_family="AluX"
repeat_region complement(26071..26107)
/rpt_family="AT_rich"
repeat_region complement(26186..26477)
/rpt_family="AluO"
repeat_region complement(26594..26757)
/rpt_family="MIR"
repeat_region 26902..27202
/rpt_family="AluSg"
repeat_region complement(28990..29296)
/rpt_family="AluO"
repeat_region 29742..29896
/rpt_family="MIR"
repeat_region complement(30043..30592)
/rpt_family="MLTIF"
repeat_region complement(31090..31367)
/rpt_family="AluX"
repeat_region 32074..32199
/rpt_family="MIR"
repeat_region 32206..32521
/rpt_family="AluSp"
repeat_region 32572..32757
/rpt_family="MIR"
repeat_region complement(32761..32838)
/rpt_family="MER2"
repeat_region complement(33407..33553)
/rpt_family="MER5B"
repeat_region complement(33575..33756)
/rpt_family="MER5A"
repeat_region complement(33771..33973)
/rpt_family="MER5A"
repeat_region 34187..34271
/rpt_family="MIR"
repeat_region 34988..35288
/rpt_family="AluX"
repeat_region 35412..35477
/rpt_family="MER5A"
repeat_region 35620..35692
/rpt_family="L1PA2"
repeat_region 36302..36575
/rpt_family="L1MB5"
repeat_region 36577..36686
/rpt_family="MIR"
repeat_region complement(36933..37007)
/rpt_family="(TGA)n"
repeat_region 37875..37999
/rpt_family="FLAM_A"
repeat_region 38476..38556
/rpt_family="PTR5"
repeat_region 38557..38849
/rpt_family="AluY"
repeat_region 38850..39247
/rpt_family="LTR12"
repeat_region complement(39625..39748)
/rpt_family="MIR"
repeat_region 40650..40674
/rpt_family="(CAG)n"
repeat_region 42135..42188
/rpt_family="(GGA)n"
repeat_region complement(43811..43983)
/rpt_family="MIR"
repeat_region 44313..44493
/rpt_family="MER5B"
repeat_region complement(44546..44620)
/rpt_family="MER5A"
repeat_region complement(45397..45480)
/rpt_family="L1MB8"
repeat_region 45317..45584
/rpt_family="MLTIF"

```

Query Match

80.8%; Score 19.4; DB 9; Length 104147;

```

Best Local Similarity 95.2%; Pred. No. 27;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CTGGATCTCAGTCATCA 22
Db 45121 CTGGATCTCAGTCATCA 45101

```

RESULT 6

AC146691

```

LOCUS 152056 bp DNA linear PRI 03-FEB-2004
DEFINITION Macaca mulatta clone ch250-214a10, complete sequence.
ACCESSION AC146691
VERSION AC146691.5 GI:41684679
KEYWORDS HTG.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Molecular Data: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
1 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Rhesus macaque BAC clone ch250-214a10
Unpublished
2 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (15-JAN-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (27-JAN-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (03-FEB-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Feb 3, 2004 this sequence version replaced gi:41349941.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----

```

REFERENCE

AUTHORS

Qin,B., Lyons,L. and Roe,B.A.

Direct Submission

Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

3 (bases 1 to 152056)

Qin,B., Lyons,L. and Roe,B.A.

Direct Submission

Submitted (15-JAN-2004) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

4 (bases 1 to 152056)

Qin,B., Lyons,L. and Roe,B.A.

Direct Submission

Submitted (27-JAN-2004) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

5 (bases 1 to 152056)

Qin,B., Lyons,L. and Roe,B.A.

Direct Submission

Submitted (03-FEB-2004) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

On Feb 3, 2004 this sequence version replaced gi:41349941.

----- Genome Center

Center: Department Of Chemistry And Biochemistry

The University Of Oklahoma

Center code:UOKNOR

Location/Qualifiers

1..152056

/organism="Macaca mulatta"

/mol_type="genomic DNA"

/db_xref="taxon:9544"

/clone="ch250-214a10"

/clone_1b="CHORI-250 Rhesus macaque BAC library"

ORIGIN

Query Match

80.8%; Score 19.4; DB 9; Length 152056;

Best Local Similarity 95.2%; Pred. No. 28;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGGATCTCAGTCATCA 24

Db 45828 GGGATCTCAGTCATCA 45848

RESULT 7

AC002324
 LOCUS 171952 bp DNA linear ROD 30-MAY-1998
 DEFINITION Mus musculus chromosome 11, clone 475_H_14, complete sequence.
 AC002324
 VERSION AC002324.1 GI:3169207
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 171952)
 Birren, B., Pasman, K., Linton, L., Nusbaum, C. and Lander, E.
 Mus musculus chromosome 11, clone 475_H_14
 Unpublished
 2 (bases 1 to 171952)
 Hawkins, T.L., Reeve, M.P., Christofferson, A., Birren, B.W.,
 Pasman, K.H., Lander, E.S., McKernan, K., Munro, C., Richardson, P.,
 Barna, N., Chang, A., Cooke, P., Daly, M.J., Devon, K., Dewar, K.,
 Forrest, C., Gage, D., Geraghty, K., Guitau, G., Hagos, B., Huang, J.,
 Jacotot, L., Lane, M., Lee, K., Mackenzie, J., Margulis, N.,
 McDermott, J., Molla, M., Moloney, N., Morrow, J., Nachman, A.,
 Naylor, J., O'Connor, T., Olotu, A., Peterson, K., Rollins, G.,
 Spencer, J., Stiliwell, U., Stone, C., Strickland, C., Sydney, K.,
 Traish, A., Wilmer, F., Zemseva, I. and Zody, M.
 Direct Submission
 Submitted (15-JUL-1997) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 171952)
 Birren, B., Pasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
 Baker, J., Baldwin, O., Barna, N., Beckert, R., Ben, U., Boatin, C.,
 Boutwell, C., Brown, A., Castle, A., Cerny, J., Cooke, P., Depayre, E.,
 Devon, K., Dewar, K., Donelan, L., Etemad, S., Ferreira, P.,
 Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S.,
 Gensheimer, S., Geraghty, K., Gilmarin, T., Grant, G., Hagos, B.,
 Harris, K., Horton, L., Howland, J.C., Hul, L., Jacotot, L., Kam, L.,
 MacDonald, P., Margulis, N., McKernan, P., McGuire, A., McKernan, K.,
 Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
 Nachman, A., Naylor, J., O'Connor, T., Pavlin, B.,
 Peterson, K., Riley, R., Roberts, D., Rossello, R., Roy, A., Shyam, R.,
 Stange-Thomann, N., Stiliwell, J., Stojanovic, N., Stone, C.,
 Strickland, C., Subramanian, A., Tortorella-Miller, I., Vasilev, H.,
 Vo, A., Wegner, A., Wang, B., Wheeler, J., Wu, Y., Ye, W.D., Zhao, J. and
 Zody, M.
 Direct Submission
 Submitted (30-MAY-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 30, 1998 this sequence version replaced gi:3142319.
 All repeats were identified using RepeatMasker: Smit, A.F.A. &
 Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html.
 Location/Qualifiers
 1. 171952
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /map="11"
 /clone="475_H_14"
 /clone_lib="Research Genetics/Cal Tech CITB-HSP-C (plates
 195-384)"
 371. 456
 repeat_region
 /rpt_family="PB1D10"
 complement(1712. .827)
 repeat_region
 /rpt_family="PB1D10"
 complement(190. .1029)
 repeat_region
 /rpt_family="PB1D9"
 complement(1309. .1451)
 repeat_region
 /rpt_family="B1_MM"
 complement(1882. .2071)
 repeat_region
 /rpt_family="URR1B"
 2523. 2563
 repeat_region
 /rpt_family="B3A"
 2536. 2717
 repeat_region

repeat_region
 /rpt_family="B3"
 complement(4356. .4383)
 repeat_region
 /rpt_family="POLY_A"
 5275. 5494
 repeat_region
 /rpt_family="B3"
 5501. 5674
 repeat_region
 /rpt_family="B3"
 6287. 6402
 repeat_region
 /rpt_family="PB1D10"
 complement(7238. .7371)
 repeat_region
 /rpt_family="L2"
 7659. 7737
 repeat_region
 /rpt_family="B1-F"
 complement(8284. .8429)
 repeat_region
 /rpt_family="B1_MM"
 8576. 8652
 repeat_region
 /rpt_family="L2"
 complement(11209. .11411)
 repeat_region
 /rpt_family="B3"
 complement(11436. .11605)
 repeat_region
 /rpt_family="B3"
 complement(11634. .11836)
 repeat_region
 /rpt_family="B2"
 complement(11861. .12013)
 repeat_region
 /rpt_family="B2"
 complement(12262. .12303)
 repeat_region
 /rpt_family="PB1D9"
 12306. 12347
 repeat_region
 /rpt_family="L2"
 complement(12573. .12596)
 repeat_region
 /rpt_family="(CCAA)n"
 complement(12639. .12762)
 repeat_region
 /rpt_family="B4"
 complement(12807. .12872)
 repeat_region
 /rpt_family="PB1D7"
 12920. 13046
 repeat_region
 /rpt_family="B1-F"
 complement(13101. .13294)
 repeat_region
 /rpt_family="B2"
 13310. 13532
 repeat_region
 /rpt_family="MER2"
 complement(13534. .13650)
 repeat_region
 /rpt_family="B1_MM"
 complement(13597. .13670)
 repeat_region
 /rpt_family="B4A"
 complement(13684. .13829)
 repeat_region
 /rpt_family="B1-F"
 complement(13804. .13837)
 repeat_region
 /rpt_family="MT2B"
 13897. 13922
 repeat_region
 /rpt_family="MT2B"
 complement(13899. .14050)
 /note="Single-stranded terminator coverage."
 13923. 14070
 repeat_region
 /rpt_family="B1_MM"
 complement(14207. .14351)
 repeat_region
 /rpt_family="B1_MM"
 complement(14958. .15399)
 repeat_region
 /rpt_family="MTC"
 16021. 16143
 repeat_region
 /rpt_family="PB1D10"
 16153. 16236
 repeat_region
 /rpt_family="PB1D7"
 complement(17105. .17224)
 repeat_region
 /rpt_family="(CA)n"
 complement(17227. .17290)
 repeat_region
 /rpt_family="(CA)n"
 complement(18858. .18900)
 repeat_region
 /rpt_family="AT_rich"
 complement(18901. .18907)
 repeat_region
 /rpt_family="B1-F"
 complement(18908. .18966)
 repeat_region
 /rpt_family="B4"


```

repeat_region      complement(18967, .19101)
                    /rpt_family="B1-F"
repeat_region      complement(19543, .19690)
                    /rpt_family="RSINB1"
repeat_region      complement(19713, .19815)
                    /rpt_family="PB1D10"
repeat_region      19950, .20032
                    /rpt_family="PB1D9"
repeat_region      21057, .21132
                    /rpt_family="PB1D7"
repeat_region      21151, .21201
                    /rpt_family="ID4"
repeat_region      21438, .21541
                    /rpt_family="B1-F"
repeat_region      complement(24457, .24557)
                    /rpt_family="PB1D9"
repeat_region      25092, .25226
                    /rpt_family="B1-MM"
repeat_region      complement(25295, .25403)
                    /rpt_family="PB1D7"
repeat_region      complement(25668, .25852)
                    /rpt_family="RSINB1"
repeat_region      complement(25985, .26117)
                    /rpt_family="B1-MM"
repeat_region      26126, .26176
                    /rpt_family="B3A"
repeat_region      complement(26292, .26359)
                    /rpt_family="ID4"
repeat_region      complement(28326, .28816)
                    /rpt_family="L1-MM"
repeat_region      29393, .29599
                    /rpt_family="B2"
repeat_region      29603, .29823
                    /rpt_family="B3"
repeat_region      29996, .30193
                    /rpt_family="B3"
repeat_region      complement(30412, .30563)
                    /rpt_family="B1-MM"
repeat_region      complement(31242, .31293)
                    /rpt_family="PB1D7"
repeat_region      31515, .31647
                    /rpt_family="B1-MM"
repeat_region      32490, .32579
                    /rpt_family="PB1D10"
repeat_region      33376, .33409
                    /note="Single-stranded terminator coverage."
repeat_region      33580, .33671
                    /rpt_family="GC rich"
repeat_region      complement(35248, .35318)
                    /rpt_family="(TA)n"
repeat_region      35664, .35809
                    /rpt_family="RSINB1"
repeat_region      36199, .36304
                    /rpt_family="PB1D9"
repeat_region      complement(36426, .36535)

Query Match      80.84; Score 19.4; DB 10; Length 171952;
Best Local Similarity 95.24; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy      4 GGGATCTCAGTGCATCCACA 24
Db      73085 GGGATCTCAGTGCATCCACA 73105

RESULT 8
AC126544
LOCUS      Homo sapiens chromosome 17, clone RP11-707023, complete sequence.
DEFINITION
AC126544
VERSION    AC126544.5 GI:24111035
KEYWORDS   Homo sapiens (human)
SOURCE

```

```

ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      1 (bases 1 to 172940)
TITLE      Birren,B., Nusbaum,C. and Lander,E.
JOURNAL      Homo sapiens chromosome 17, clone RP11-707023
REFERENCE      2 (bases 1 to 172940)
AUTHORS      Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
              Barta,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
              Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
              Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
              Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,
              Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
              Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
              Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
              McCarthy,M., Meldrim,J., Meneus,L., Milnova,T., Miengwa,V.,
              Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
              O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
              Punnkhang,P., Pierre,N., Raymond,C., Retta,R., Seaman,S., Severy,P.,
              Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
              Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
              Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
              Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
              Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE      Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome
JOURNAL      Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE      3 (bases 1 to 172940)
AUTHORS      Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
              Barta,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
              Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
              Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
              Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,
              Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
              Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
              Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
              McCarthy,M., Meldrim,J., Meneus,L., Milnova,T., Miengwa,V.,
              Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
              O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
              Punnkhang,P., Pierre,N., Raymond,C., Retta,R., Seaman,S., Severy,P.,
              Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
              Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
              Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
              Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
              Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE      Submitted (10-OCT-2002) Whitehead Institute/MIT Center for Genome
JOURNAL      Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE      4 (bases 1 to 172940)
AUTHORS      Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
              Barta,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
              Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
              Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
              Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,
              Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Hages,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
              Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
              Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Milnova,T.,
              Miengwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
              Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
              Peterson,K., Punnkhang,P., Pierre,N., Raymond,C., Retta,R.,
              Seaman,S., Severy,P., Roman,J., Roy,A., Schauer,S., Schupack,R.,
              Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
              Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
              Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
              Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE      Submitted (18-OCT-2002) Whitehead Institute/MIT Center for Genome
JOURNAL      Research, 320 Charles Street, Cambridge, MA 02141, USA

```

COMMENT

On Oct 18, 2002 this sequence version replaced gi:23683037.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: I27758

Center clone name: 707_O_23

Only the first 172,9 kilobases of this clone are being submitted.
The remainder overlaps accession number AC106030 [WIGR project
I24005].

FEATURES

source

1. .172940
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-707023"
/clone_lib="RPCI-11 Human Male BAC"
complement(75..230)
/rpt_family="MIR"
repeat_region
1065..1341
/rpt_family="AluSc"
repeat_region
1344..1625
/rpt_family="AluSc"
1505..1520
/note="<30 qual SNGL region"
1531..1536
/note="<30 qual SNGL region"
1545..1554
/note="<30 qual SNGL region"
1558..1563
/note="<30 qual SNGL region"
1582..1589
/note="<30 qual SNGL region"
1849..2142
/rpt_family="AluSg"
repeat_region
2198..2493
/rpt_family="AluSc"
complement(2556..2597)
repeat_region
/rpt_family="5S"
2719..3001
repeat_region
/rpt_family="AluSc"
3002..3075
repeat_region
/rpt_family="TTAA)n"
4446..4600
repeat_region
/rpt_family="L2"
complement(4691..4987)
repeat_region
5177..5197
/rpt_family="CAAAA)n"
5318..5610
repeat_region
/rpt_family="AluSg"
6449..6732
repeat_region
/rpt_family="AluSc"
6745..6859
repeat_region
/rpt_family="FLAM_A"
complement(6872..7027)
repeat_region
/rpt_family="AluSg"
complement(7028..7092)
repeat_region
/rpt_family="MERSA"
7432..7565
repeat_region
/rpt_family="MERSA"
8173..8214
repeat_region
/rpt_family="L2"
complement(8589..8881)
repeat_region

repeat_region /rpt_family="AluSc"
complement(8882..8997)
repeat_region /rpt_family="MIR"
complement(9021..9375)
repeat_region /rpt_family="L1MDa"
complement(9594..9668)
repeat_region /rpt_family="L2"
9672..9956
repeat_region /rpt_family="AluSc"
10134..10281
repeat_region /rpt_family="(TA)n"
10282..10310
repeat_region /rpt_family="(CA)n"
10311..10704
repeat_region /rpt_family="L1MB1"
10707..10823
repeat_region /rpt_family="FLAM_C"
10948..10965
repeat_region /rpt_family="L1MB1"
complement(10966..11261)
repeat_region /rpt_family="AluSc"
complement(11270..11570)
repeat_region /rpt_family="AluSc"
11571..12181
repeat_region /rpt_family="L1MB1"
12205..12506
repeat_region /rpt_family="AluSc"
12516..12646
repeat_region /rpt_family="FLAM_C"
12653..13174
repeat_region /rpt_family="L1MA4"
13312..13506
repeat_region /rpt_family="MERS8B"
13515..13735
repeat_region /rpt_family="L1MB1"
13739..14044
repeat_region /rpt_family="AluDb"
14047..14340
repeat_region /rpt_family="AluY"
14366..14419
repeat_region /rpt_family="L1MB1"
14420..14718
repeat_region /rpt_family="AluSc"
14719..14884
repeat_region /rpt_family="L1MB1"
14885..15188
repeat_region /rpt_family="AluY"
15189..15551
repeat_region

Query Match 80.8%; Score 19.4; DB 9; Length 172940;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGGATCTCAGTCATCCAA 22

Db 109700 CTGGATCTCAGTCATCCAA 109720

RESULT 9

AC025584/c 192550 bp DNA linear HTG 06-SEP-2000
LOCUS AC025584
DEFINITION Mus musculus clone RP23-263C10, WORKING DRAFT SEQUENCE, 20
unordered pieces.

ACCESSION AC025584
VERSION AC025584.3 GI:9972308
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 192550)
McCombie,W.R., Baker,J.P., Bahret,A., Bal,H., Dedhia,N.N., de la
Bastide,W., Huang,E.N., King,L., Kirchoff,K.A., Miller,B.,

TITLE
JOURNAL
AUTHORS
JOURNAL

REFERENCE
TITLE
JOURNAL

COMMENT

Naecimento, L.U., O'Shaughnessy, A.L., Preston, R.R., Rodriguez, M.A.,
Shar, R.S., Shakhshar, M., Spiegel, L.A., Toth, K. and Vill, M.D.

Mouse Genomic Sequence

Unpublished

2 (bases 1 to 192550)

McCombie, W.R.

Submitted (11-MAR-2000)

Center: Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA

On Sep 6, 2000 this sequence version replaced gi:8439865.

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor

Laboratory

Center code: CSHL

Web site: <http://www.cshl.org/genseq>

Contact: mccombie@cshl.org

Project Information

Center project name: RP23-263C10

Center clone name: RP23-263C10

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 42099: contig of 42099 bp in length
42100 42499: gap of unknown length
42500 59653: contig of 17154 bp in length
59654 60053: gap of unknown length
60054 76545: contig of 16492 bp in length
76546 76945: gap of unknown length
76946 90821: contig of 13876 bp in length
90822 91221: gap of unknown length
91222 105073: contig of 13852 bp in length
105074 105473: gap of unknown length
105474 117737: contig of 12264 bp in length
117738 118137: gap of unknown length
118138 128984: contig of 10847 bp in length
128985 129384: gap of unknown length
129385 138289: contig of 8905 bp in length
138290 138689: gap of unknown length
138690 147015: contig of 8326 bp in length
147016 147415: gap of unknown length
147416 153692: contig of 6277 bp in length
153693 154092: gap of unknown length
154093 159704: contig of 5612 bp in length
159705 160104: gap of unknown length
160105 165289: contig of 5185 bp in length
165290 165689: gap of unknown length
165690 170332: contig of 4643 bp in length
170333 170732: gap of unknown length
170733 175202: contig of 4470 bp in length
175203 175602: gap of unknown length
175603 178792: contig of 3180 bp in length
178793 179192: gap of unknown length
179193 182135: contig of 2943 bp in length
182136 182535: gap of unknown length
182536 185383: contig of 2848 bp in length
185384 185783: gap of unknown length
185784 188478: contig of 2635 bp in length
188479 188878: gap of unknown length
188879 191019: contig of 2141 bp in length
191020 191418: gap of unknown length
191419 192550: contig of 1132 bp in length.

FEATURES
SOURCE

1. 192550
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN

/clone="RP23-263C10"

Query Match 80.8%; Score 19.4; DB 2; Length 192550;

Best Local Similarity 95.2%; Pred. No. 28;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

4 GGGATCTCAGTCATCCACA 24

Db 191770 GGGATCTCAGTCATCCACA 191750

RESULT 10

AL591177/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:16304973.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Bm, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-399H5 is
from the RPCI-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pPac3.6.

FEATURES

source

1. 193004
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-399H5"
/clone_11b="RPCI-23"

ORIGIN

Query Match 80.8%; Score 19.4; DB 10; Length 193004;

Best Local Similarity 95.2%; Pred. No. 28;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

4 GGGATCTCAGTCATCCACA 24

Db 11904 GGGATCTCAGTCATCCACA 11884

RESULT 11


```

/note="assembly_name:Contig38
clone_end:SP6
vector_side:right"
misc_feature 146329..164594
/note="assembly_name:Contig39"
misc_feature 164695..198388
/note="assembly_name:Contig40"

Query Match 80.8%; Score 19.4; DB 2; Length 198388;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 183265 CTGGATCTCAGTCATCCAA 183245
2 CTGGATCTCAGTCATCCAA 22
|||||
AC126546 206578 bp DNA 11near HTG 03-OCT-2002
LOCUS Homo sapiens chromosome 15 clone RP11-733021 map 15, WORKING DRAFT
DEFINITION AC126546
SEQUENCE, 16 unordered pieces.
AC126546
AC126546.2 GI:23477883
VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 206578)
Birken, B., Nuebaum, C. and Lander, E.
Homo sapiens chromosome 15, clone RP11-733021
Unpublished
2 (bases 1 to 206578)
Birken, B., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Gardyn, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horizon, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Kartas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Menes, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 206578)
Birken, B., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Gardyn, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horizon, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Kartas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Menes, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

```

TITLE JOURNAL COMMENT

```

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (03-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 3, 2002 this sequence version replaced GI:21700691.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27774
Center clone name: 733-O_21
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960721
Consensus quality: 20135 bases at least Q40
Consensus quality: 203359 bases at least Q30
Consensus quality: 204367 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 205078; sum-of-contigs
Quality coverage: 15.0 in Q20 bases; agarose-fp
Quality coverage: 13.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1796: contig of 1796 bp in length
1896: gap of 100 bp
1897 2574: contig of 678 bp in length
2575 2674: gap of 100 bp
2675 3678: contig of 1004 bp in length
3679 3778: gap of 100 bp
3779 4472: contig of 694 bp in length
4473 4572: gap of 100 bp
4573 5587: contig of 1015 bp in length
5588 5687: gap of 100 bp
5688 6417: contig of 730 bp in length
6418 6517: gap of 100 bp
6518 7278: contig of 761 bp in length
7279 7378: gap of 100 bp
7379 8745: contig of 1367 bp in length
8746 8845: gap of 100 bp
8846 9956: contig of 1111 bp in length
9957 10056: gap of 100 bp
10057 11931: contig of 3875 bp in length
11932 14031: gap of 100 bp
14032 22254: contig of 8223 bp in length
22255 22354: gap of 100 bp
22355 29263: contig of 6909 bp in length
29264 29363: gap of 100 bp
29364 64453: contig of 35090 bp in length
64454 64553: gap of 100 bp
64554 91996: contig of 27443 bp in length
91997 92096: gap of 100 bp
92097 124248: contig of 33152 bp in length
124249 124348: gap of 100 bp
124349 206578: contig of 82230 bp in length.
-----
FEATURES
source
1..206578
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

/misc_feature /chromosome="15"
/misc_feature /map="15"
/misc_feature /clone="RP11-733021"
/misc_feature /clone.lib="RP11-11 Human Male BAC"
1.1796
/misc_feature /note="assembly_fragment"
/misc_feature /note_end:SP6
/misc_feature /note_end:left"
1897.2574
/misc_feature /note="assembly_fragment"
/misc_feature 2675.3678
/misc_feature /note="assembly_fragment"
/misc_feature 3779.4472
/misc_feature /note="assembly_fragment"
/misc_feature 4573.5587
/misc_feature /note="assembly_fragment"
/misc_feature 5688.6417
/misc_feature /note="assembly_fragment"
/misc_feature 6518.7278
/misc_feature /note="assembly_fragment"
/misc_feature 7379.8745
/misc_feature /note="assembly_fragment"
/misc_feature 8846.9956
/misc_feature /note="assembly_fragment"
/misc_feature 10057.13931
/misc_feature /note="assembly_fragment"
/misc_feature 14032.22254
/misc_feature /note="assembly_fragment"
/misc_feature 22355.22263
/misc_feature /note="assembly_fragment"
/misc_feature 29364.64453
/misc_feature /note="assembly_fragment"
/misc_feature 64554.91996
/misc_feature /note="assembly_fragment"
/misc_feature 92097.124248
/misc_feature /note="assembly_fragment"
/misc_feature 124349.206578
/misc_feature /note="assembly_fragment"

```

ORIGIN

```

Query Match 80.8%; Score 19.4; DB 2; Length 206578;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy 2 CTGGGATCTCAGTCATCCAA 22
Db 156371 CTGGGATCTCAGTCATCCAA 156351

```

RESULT 13

```

AC107969 57662 bp DNA linear HTG 24-JAN-2002
LOCUS Homo sapiens chromosome 11 clone CTD-2028015 map 11, LOW-PASS
DEFINITION
SEQUENCE SAMPLING.
AC107969
AC107969.1 GI:18308697
VERSION HTG, HTGS PHASEO.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 11, clone CTD-2028015
REFERENCE
AUTHORS Unpublished
2 (bases 1 to 57662)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barin,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarero,J., Campopiano,A., Chang,J., Chazaro,B.,
Chen,P., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,N.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

```

TITLE

JOURNAL

COMMENT

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karakas,A., Kells,C., Lacroque,K., Lamazares,R., Lander,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Margis,N., Matthews,C., McCarthy,M., McKwan,P., McKernan,K., Meldrum,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Ries,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severi,P., Spencer,B., Stange-Thoman,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tefaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: l24518

Center clone name: 2028_O15

```

***** NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
718 717: contig of 717 bp in length
818 817: gap of 100 bp
818 1525: contig of 708 bp in length
1526 1625: gap of 100 bp
1626 2265: contig of 640 bp in length
2266 2365: gap of 100 bp
2366 3050: contig of 685 bp in length
3051 3150: gap of 100 bp
3151 3845: contig of 695 bp in length
3846 3945: gap of 100 bp
3946 4577: contig of 632 bp in length
4578 4677: gap of 100 bp
4678 5368: contig of 692 bp in length
5370 5469: gap of 100 bp
5470 6151: contig of 682 bp in length
6152 6251: gap of 100 bp
6252 6955: contig of 704 bp in length
6956 7055: gap of 100 bp
7056 7804: contig of 749 bp in length
7805 7904: gap of 100 bp
7905 8603: contig of 699 bp in length
8604 8703: gap of 100 bp
8704 9392: contig of 689 bp in length
9393 9492: gap of 100 bp
9493 10135: contig of 663 bp in length
10136 10255: gap of 100 bp
10256 10948: contig of 693 bp in length
10949 11048: gap of 100 bp
11049 11603: contig of 555 bp in length
11604 11703: gap of 100 bp
11704 12386: contig of 683 bp in length
12387 12486: gap of 100 bp

```

```

12487 13187: contig of 701 bp in length
13188 13287: gap of 100 bp
13288 13988: contig of 701 bp in length
13989 14088: gap of 100 bp
14089 14798: contig of 707 bp in length
14799 14898: gap of 100 bp
14899 15532: contig of 637 bp in length
15533 15632: gap of 100 bp
15633 16347: contig of 715 bp in length
16348 16447: gap of 100 bp
16448 17155: contig of 708 bp in length
17156 17255: gap of 100 bp
17256 17961: contig of 706 bp in length
17962 18061: gap of 100 bp
18062 18723: contig of 662 bp in length
18724 18823: gap of 100 bp
18824 19476: contig of 653 bp in length
19477 19576: gap of 100 bp
19577 20274: contig of 698 bp in length
20275 20374: gap of 100 bp
20375 21074: contig of 700 bp in length
21075 21174: gap of 100 bp
21175 21867: contig of 693 bp in length
21868 21967: gap of 100 bp
21968 22680: contig of 713 bp in length
22681 22780: gap of 100 bp
22781 23495: contig of 715 bp in length
23496 23595: gap of 100 bp
23596 24227: contig of 632 bp in length
24228 24327: gap of 100 bp
24328 25032: contig of 705 bp in length
25033 25132: gap of 100 bp
25133 25838: contig of 706 bp in length
25839 25938: gap of 100 bp
25939 26651: contig of 713 bp in length
26652 26751: gap of 100 bp
26752 27431: contig of 680 bp in length
27432 27531: gap of 100 bp
27533 28332: contig of 701 bp in length
28333 28332: gap of 100 bp
28333 29032: contig of 700 bp in length
29033 29132: gap of 100 bp
29133 29638: contig of 706 bp in length
29639 29938: gap of 100 bp
29939 30653: contig of 715 bp in length
30654 30753: gap of 100 bp
30754 31466: contig of 713 bp in length
31467 31566: gap of 100 bp
31567 32272: contig of 706 bp in length
32273 32372: gap of 100 bp
32373 33009: contig of 637 bp in length
33010 33109: gap of 100 bp
33110 33627: contig of 718 bp in length
33628 33927: gap of 100 bp
33928 34646: contig of 719 bp in length
34647 34746: gap of 100 bp
34747 35451: contig of 705 bp in length
35452 35551: gap of 100 bp
35552 36244: contig of 693 bp in length
36245 36344: gap of 100 bp
36345 36993: contig of 649 bp in length
36994 37093: gap of 100 bp
37094 37777: contig of 684 bp in length
37778 37877: gap of 100 bp
37878 38587: contig of 710 bp in length
38588 38687: gap of 100 bp
38688 39375: contig of 688 bp in length
39376 39475: gap of 100 bp
39476 40179: contig of 704 bp in length
40180 40279: gap of 100 bp
40280 40986: contig of 707 bp in length
40987 41086: gap of 100 bp
41087 41749: contig of 663 bp in length

```

```

41750 41849: gap of 100 bp
41850 42551: contig of 702 bp in length
42552 43363: contig of 712 bp in length
43364 43463: gap of 100 bp
43464 44144: contig of 681 bp in length
44145 44244: gap of 100 bp
44245 44939: contig of 695 bp in length
44940 45039: gap of 100 bp
45040 45740: contig of 701 bp in length
45741 45840: gap of 100 bp
45841 46513: contig of 673 bp in length
46514 46613: gap of 100 bp
46614 47307: contig of 694 bp in length
47308 47407: gap of 100 bp
47408 48097: contig of 690 bp in length
48098 48197: gap of 100 bp
48198 48884: contig of 687 bp in length
48885 48984: gap of 100 bp
48985 49700: contig of 716 bp in length
49701 49800: gap of 100 bp
49801 50484: contig of 684 bp in length
50485 50584: gap of 100 bp
50585 51288: contig of 704 bp in length
51289 51388: gap of 100 bp
51389 52107: contig of 719 bp in length
52108 52207: gap of 100 bp
52208 52809: contig of 702 bp in length
52810 53009: gap of 100 bp
53010 53665: contig of 656 bp in length
53666 53765: gap of 100 bp
53766 54444: contig of 679 bp in length

```

Query Match
Best Local Similarity 80.0%; Score 19.2; DB 2; Length 57662;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCTGGATCTGATGCATCAACA 24
Db 1359 TCTGGATTTCATGCATCA 1382

RESULT 14
AL356577
LOCUS Homo sapiens chromosome 11 clone RPI-6104.
DEFINITION Homo sapiens chromosome 11 clone RPI-6104.
ACCESSION AL356577
VERSION AL356577.5 GI:11991384
KEYWORDS HTG; HTGS PHASE2; HTGS_CANCELED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Martin, S.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Dec 24, 2000 this sequence version replaced gi:11931890.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: d66104
----- Summary Statistics
Sequencing program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 77% of reads
Chemistry: Dye-terminator ET-amersham; 22% of reads
Consensus quality: 8455 bases at least Q40

Consensus quality: 84806 bases at least Q30
Consensus quality: 84862 bases at least Q20
Insert size: 85378, sum-of-configs
Insert size: 98007, 13.5% error; agarose-fp
Quality coverage: 8.00x in Q20 bases; sum-of-configs
coverage: 7.06x in Q20 bases; agarose-fp

misc_Feature	1. .85378	/note="assembly_Fragment:01465"			
ORIGIN					
Query Match	80.0%	Score 19.2;	DB 2;	Length 65378;	
Best Local Similarity	87.5%	Pred. No.3;			
Matches	21; Conservative	0; Mismatches	3;	Indels	0; Gaps
					0

RESULT 15	
HSJD247C2	HSJD247C2 98360 bp DNA linear PRI 03-AUG-2001
LOCUS	
DEFINITION	Human DNA sequence from clone RPI-247C2 on chromosome 1p13
ACCESSION	Contains STSs and GSSs, complete sequence.
VERSION	AL049713
KEYWORDS	AL049713.20 GI:10334639
SOURCE	HTG.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	1 (bases 1 to 98360)
TITLE	Wallis, J.
JOURNAL	Direct Submission
	Submitted (02-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,

assembly was confirmed by restriction digest. Rpl-247C2 is from the library RPlC1-1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2. (GenBank accession number: AF031454)

URES	Location/Qualifiers
source	1. .98360
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="P13"
	/clone="RP1-247C2"
	/clone_1kb="RPC1-1"
misc_feature	complement(627. .983)
	/note="match: STS: Em:HS198YB10"
repeat_region	804. .829
	/note="13 copies 2 mer gt 100% conserved"
repeat_region	3773. .3810
	/note="L1PA11 repeat: matches 2154. .2191 of consensus"
repeat_region	3956. .3992
	/note="L1PA11 repeat: matches 2180. .2215 of consensus"
repeat_region	4265. .5286
	/note="L1PA11 repeat: matches 2215. .3216 of consensus"
repeat_region	5586. .8453
	/note="L1PA11 repeat: matches 3216. .6162 of consensus"
repeat_region	10697. .10906
	/note="L1MEC repeat: matches 2142. .1979 of consensus"
repeat_region	11202. .11493
	/note="L1MEC repeat: matches 1978. .2279 of consensus"
repeat_region	13794. .13857
	/note="13 copies 2 mer cc 68% conserved"
repeat_region	14618. .15078
	/note="L1R2 repeat: matches 1. .449 of consensus"
misc_feature	complement(14642. .15322)
	/note="match: GSS: Em:B04935"
misc_feature	15910. .15911
	/note="single clone region. Assembly confirmed by restriction digest data."
repeat_region	15930. .15955
	/note="13 copies 2 mer tg 92% conserved"
repeat_region	16944. .17077
	/note="L1R1A2 repeat: matches 1. .92 of consensus"
repeat_region	17380. .17654
	/note="L1R1A2 repeat: matches 92. .374 of consensus"
misc_feature	complement(17710. .18372)
	/note="match: GSS: Em:AQ485260"
misc_feature	18444. .18750
	/note="match: GSS: Em:AQ545989"
misc_feature	complement(20477. .21170)
	/note="match: GSS: Em:AQ486644"
misc_feature	complement(20761. .21177)
	/note="match: GSS: Em:AQ698825"
misc_feature	complement(25626. .26158)
	/note="match: GSS: Em:AQ505438"
misc_feature	25991. .26104
	/note="match: GSS: Em:A2393320"
misc_feature	26106. .26561
	/note="match: GSS: Em:A2393320"
repeat_region	29550. .29636
	/note="L1MA8 repeat: matches 6175. .6268 of consensus"
repeat_region	33434. .33473
	/note="L1R29 repeat: matches 56. .98 of consensus"
misc_feature	complement(34597. .34987)
	/note="match: GSS: Em:AQ877394"
repeat_region	35999. .36384
	/note="L1PB3 repeat: matches 5743. .6147 of consensus"
misc_feature	37497. .37542
	/note="23 copies 2 mer tg 100% conserved"
misc_feature	38577. .38776
	/note="match: STS: Em:G45608"
misc_feature	38617. .38781


```

misc_feature      /note="match: STS: Em:A0048503"
38620..38722
/note="match: STS: Em:A0046429"
misc_feature      complement(38660..38777)
/note="match: GSS: Em:A2398168"
38685..38746
/note="31 copies 2 mer gc 91% conserved"
repeat_region     38956..39215
/note="Charlieia repeat: matches 28..281 of consensus"
40351..40840
/note="Charlieia repeat: matches 277..769 of consensus"
42053..42372
/note="Charlieia repeat: matches 778..1137 of consensus"
44071..44491
/note="match: GSS: Em:A0112832"
misc_feature      complement(45407..45624)
/note="match: GSS: Em:A0247033"
53815..53904
/note="Charlie2 repeat: matches 3272..3370 of consensus"
54009..54235
/note="Charlie2 repeat: matches 2888..3117 of consensus"
56236..57009
/note="LIPa2 repeat: matches 1..776 of consensus"
57005..62248
/note="LIPa2 repeat: matches 900..6144 of consensus"
63508..63541
/note="17 copies 2 mer aa 82% conserved"
63586..63668
/note="LIPD repeat: matches 738..819 of consensus"
63917..64021
/note="LIPD repeat: matches 984..1089 of consensus"
64543
/note="match: GSS: Em:A085350"
complement(66096..66577)
/note="match: GSS: Em:A0611651"
69990..70342
/note="match: STS: Em:G21603"
complement(76094..76635)
/note="match: GSS: Em:A0277578"
76636
complement(76184..76636)
/note="match: GSS: Em:A0812553"
complement(76460..76628)
/note="match: GSS: Em:A0140871"
complement(88581..89052)
/note="match: GSS: Em:A0201907"
89057..89594
/note="match: GSS: Em:A0543464"
91625..91990
/note="THE1C repeat: matches 1..371 of consensus"
92079..92237
/note="LIPa7 repeat: matches 5949..6125 of consensus"

```

ORIGIN

```

Query Match      80.0%; Score 19.2; DB 9; Length 98360;
Best Local Similarity 87.5%; Pred. No. 35;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Oy      1 TCTGGGATCTCAGTCATCAACA 24
          |||||
Db      36062 TCTGGGATTTAGTCATCAACA 36085

```

Search completed: October 15, 2005, 01:33:41
Job time : 104.037 secs

This Page Blank (uspto)